Understanding Segregation Change: Methods and Applications

Benjamin Elbers

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

Abstract

Understanding Segregation Change: Methods and Applications

Benjamin Elbers

Racial residential and school segregation, while having declined in recent decades, are still pervasive in U.S. metropolitan areas. Given the consequences of segregation for individual life outcomes and its role in exacerbating inequalities in the U.S., it is of major importance to better understand the processes that shape segregation. The goal of this dissertation is to develop methods that allow us to better understand which social processes are producing increases and declines in segregation.

The dissertation consists of five substantive chapters. In chapters two and four, I develop two decompositions methods that allow the decomposition of changes in segregation. The first decomposition method focuses on providing a mechanical solution to the problem of "margin dependency." Unlike alternative methods, this decomposition does not attempt to "purge" the index from its margin dependency, but instead quantifies how much of a given change in segregation is due to changes in the margins, compared to structural changes. Arguably, this method provides more information about changes in segregation than a simple trend analysis. The fourth chapter introduces a more flexible method of decomposition, which allows the researcher to specify decompositions that are guided by theoretical considerations. This decomposition method is based on the Shapley value, originally developed in game theory. This

chapter also shows that the Shapley value decomposition has many applications outside of segregation studies.

The remaining substantive chapters are applications demonstrating the usefulness of these decompositions to understand changes in segregation. The third chapter applies the marginal-structural decomposition to a topic usually not considered in segregation analysis: the study of school-to-work linkages. This coauthored paper compares the skill-formation systems of France and Germany. Stratification research has often made a distinction between two ideal-types: "qualificational spaces," exemplified by Germany with a focus on vocational education, and "organizational spaces," exemplified by France with a focus on general education. Most studies that investigated this distinction did so by focusing only on the size of the vocational sector, not on whether graduates with a vocational degree actually link strongly to the labor market. Moreover, these studies often studied male workers only, ignoring potential gender differences in how school-to-work linkages are established. Our approach is instead to map the change in education-occupation linkage in France and Germany between 1970 and 2010, using the marginal-structural decomposition to distinguish between changes in rates (marginal changes) and changes in the structure of school-to-work linkages (structural changes). Surprisingly, we find that the German vocational system in 1970 was not, on average, substantially more efficient in allocating graduates to specific occupations than the French system. This finding is a major departure from earlier results, and it shows that the differences between 1970's France and Germany, on which the qualificational-organizational distinction is based, are smaller than previously assumed. Partly, this is due to the fact that the female labor force was omitted from earlier analyses. We thus show that ignoring the female workforce has consequences for today's conception of skill formation systems, particularly because a large share of educational expansion is caused by an increase in female enrollment in (higher) education.

In the remaining two chapters, I apply the Shapley decomposition strategy to two long-standing interests of U.S. sociology: racial residential and racial school segregation. The fifth chapter, on racial residential segregation, studies changes in segregation from 1990-2010. This paper engages with a prominent concept in segregation studies, the idea of micro and macro segregation. Micro segregation refers to the small-scale neighborhood segregation within cities and suburbs, while macro segregation refers to segregation between larger geographical areas, such as cities, suburbs, and school districts. The paper first shows that, contrary to other results in literature, while micro segregation decreased, macro segregation remained at similar levels. Second, the paper shows that declines in segregation are almost exclusively caused by the Black and Hispanic populations, which have increasingly moved to majority-White areas. The sixth chapter studies changes in between-district school segregation from 2009 to 2016, studying both the relationships between school district racial composition and school district performance, as quantified by average test scores. Also in this later period, declines in segregation are mostly driven by the Black and Hispanic populations. Additionally, the decomposition by school district performance shows that families of all racial groups move from badly-performing school districts to better-performing districts.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Throughout my years as a PhD student, I received generous support from several grants and fellowships. The first to mention is the Fulbright Fellowship, which allowed me to come to New York and spend a year at the CUNY Graduate Center. I soon realized that a year in New York passes far too quickly, and, with Columbia's support, I was able to stay for another six. In the final stages of my PhD, I was supported by a grant from the American Educational Research Association which receives funds for its AERA Grants Program from the National Science Foundation under NSF award NSF-DRL 1749275, and by a grant from the Horowitz Foundation for Social Policy. Opinions reflect those of the author and do not necessarily reflect those of AERA, NSF, or the Horowitz Foundation.

I have many others to thank. In the sociology department, a special thanks goes out to Greer and Joanna, who graciously shared information on various statistics courses to attend or to avoid. These classes were much more fun with Siyeona, although, unfortunately, we had to attend most of them over Zoom. Together with Jon and Martin, I enjoyed both our squash games and, later in the evening, a couple of beers in the pub. A special thanks goes to Martin for introducing me to some of New York's culinary highlights, and for always being there to discuss the hard questions.

The incomparable Zahava, of course, deserves a special mention. Countless Shabbat dinners provided much-needed emotional support after tough weeks of data cleaning, writing, or teaching. Your down-to-earth advice, your generosity, and your optimism are an inspiration. I

cannot imagine a life in New York without you. The same goes for my amazing support group from CUNY: Thank you, Anya, Isa, and Sebastián, for road trips, movie nights, vacations, many conversations over kasha, tacos, or arepas, and so much more.

Thanks to my committee members, Tom, Yao, Shamus, Heike, and Thijs, for their feedback, encouragement, and support! Tom deserves a special mention, as he started to integrate me into the research process from day one, was always available for meetings, both short and long, and graciously opened up his network of researchers and collaborators to me. I would not be the researcher I am today without Tom's encouragement and trust. Thank you!

Lastly, I am grateful to my friends and family in the "old country," who, despite my absconding to New York, were always there for me. My final, and biggest, thanks go to Jana, who made sure that, whenever I returned from New York, I could join her in some European port city, be it Copenhagen, Hamburg, or Antwerp.

Chapter 1: Introduction

1.1 Overview

The papers collected in this dissertation are all concerned with a specific class of segregation measures. These *entropy-based* or *information-theoretic* segregation measures were first introduced into the social sciences by the Dutch econometrician Henri Theil (1967, 1971, 1972). These measures have attractive properties that make them especially useful for studying segregation problems. In the social sciences, the interest in studying a segregation problem usually does not lie only in computing a single segregation index, but also in comparing trends in segregation over time or across geographical areas. The question of how to study changes in segregation has usually received attention in the literature only due to the issue of "margin dependency." This refers to the problem that the value of an index can change mechanically when the distributions of the groups (for instance, genders or racial groups) or the units (for instance, occupations) changes. Such mechanical changes can make comparisons across time complicated, as the change in segregation that is due to changes in the marginal distributions is usually regarded as less significant compared to changes in segregation that are due to more fundamental causes.

While the focus on margin dependency is an important methodological issue—one that will be discussed in detail in the following chapter—it does not help much with understanding how changes in segregation are actually produced. For instance, studies of racial residential segregation have shown that, even after accounting for changes in the marginal distributions, Black-White segregation has declined since the 1970s. While an important fact, this trend alone tells us nothing about the reasons for this decline. As will be argued in chapters five and six, attempts to "explain" changes in segregation through the use of regression models have not furthered our understanding regarding the drivers of segregation change. In chapters two and four, I develop methods to more directly understand the drivers of segregation change through decomposition procedures. This is the main contribution of this dissertation.

The dissertation consists of six chapters, including this introduction. The following chapter is methodological, and focuses on providing a mechanical solution to the problem of understanding changes in segregation. For any study of segregation trends, it is always possible to apply this decomposition, namely into *structural* and *marginal* components. This chapter hence provides a solution to the problem of margin dependency that does not attempt to "purge" the index from its margin dependency, but instead quantifies how much of a given change in segregation is due to changes in the margins. Arguably, this method provides more information about changes in segregation than a simple trend analysis. This chapter also discusses the useful *local decomposition* property of entropy-based segregation indices, which can be used to further pinpoint the sources of change.

The third chapter applies the *marginal-structural* decomposition to a topic usually not considered in segregation analysis: the study of school-to-work linkages. This coauthored paper compares the skill-formation systems of France and Germany. Stratification research has often made a distinction between two ideal-types: "qualificational spaces," exemplified by Germany with a focus on vocational education, and "organizational spaces," exemplified by France with a focus on general education. Most studies that investigated this distinction did so by focusing only on the *size* of the vocational sector, not on whether graduates with a vocational degree actually link strongly to the labor market. Moreover, these studies often focused on male workers only, ignoring potential gender differences in how school-to-work linkages are established. Our approach is instead to map the change in education-occupation linkage in France and Germany between 1970 and 2010, using the marginal-structural decomposition to distinguish between changes in rates (marginal changes) and changes in the structure of school-to-work linkages (structural changes). Surprisingly, we find that the German vocational system in 1970 was not, on average, substantially more efficient in allocating graduates to specific occupations than the French system. This finding is a major departure from earlier results, and it shows that the differences between 1970's France and Germany, on which the qualificational-organizational distinction is based, are smaller than previously assumed. Partly, this is due to the fact that the female labor force was omitted from earlier analyses. We thus show that ignoring the female workforce has consequences for today's conception of skill formation systems, particularly because a large share of educational expansion can be attributed to the increase in female enrollment in (higher) education.

The remaining three chapters form another set of connected papers. The fourth chapter introduces a more flexible method of decomposition, which allows me to specify decompositions that are guided by theoretical considerations. This decomposition method is based on the Shapley value, developed in game theory. The Shapley value decomposition has many applications outside of segregation studies, which, however, are not considered here. The remaining two chapters apply the Shapley decomposition strategy to two long-standing interests of U.S. sociology: racial residential and racial school segregation. The fifth chapter, on racial residential segregation, studies changes in segregation from 1990 to 2010. This paper engages with a prominent concept in segregation studies, the idea of micro and macro segregation. Micro segregation refers to the small-scale neighborhood segregation *within* cities and suburbs, while macro segregation refers to segregation *between* larger geographical areas, such as cities, suburbs, and school districts. Contrary to other results in the literature, the paper first shows that, while micro segregation decreased, macro segregation remained at similar levels. Second, the paper shows that declines in segregation are almost exclusively caused by the Black and Hispanic populations, which have increasingly moved to majority-White areas. The sixth chapter studies changes in between-district school segregation from 2009 to 2016, studying both the relationships between school district racial composition and school district performance, as quantified by average test scores. Also in this later period, declines in segregation are mostly driven by the Black and Hispanic populations. Additionally, the decomposition by school district performance shows that families of all racial groups move from badly-performing school districts to better-performing districts.

All procedures discussed in this dissertation have been implemented in an R package. The package includes the calculation of the M and H indices, as well as the decomposition procedures developed in chapter 2 and 4. Standard errors can be bootstrapped for both the index calculation and the decomposition procedure. A short introduction on how to use the package can be found in the Appendix.

1.2 Connection to other literatures

For sociologists, the study of segregation is a key problem in categorical data analysis. Despite this fact, connections between segregation indices and other statistical literatures on categorical data analysis are rarely drawn. For instance, a standard textbook on categorical data analysis, by Agresti (2013), does not contain a single mention of the term "segregation."[1] Similarly, the methodological work on segregation indices does not connect to the statistical literature on categorical data analysis, and only rarely connects to neighboring literatures that analyze similar problems, such as the social mobility literature.

To show these connections, a basic definition of the segregation indices used in this dissertation is required. Any segregation index summarizes the dependency structure of a contingency table. Throughout this dissertation, I will use the following notation: Let $\mathbf{T}$ be a matrix with U rows (often, spatial units or occupations) and G columns (often, racial or gender groups). In most segregation problems, $U \gg G$. Let the entries of $\mathbf{T}$ be t_{ug}, as in Table 1.1, and let t be the total population of $\mathbf{T}$, i.e. $t = \sum_{u=1}^{U} \sum_{g=1}^{G} t_{ug}$. The estimated joint probability of being in unit u and group g is $p_{ug} = t_{ug}/t$. Also define $p_{u\cdot} = \sum_{g=1}^{G} t_{ug}/t$ and $p_{\cdot g} = \sum_{u=1}^{U} t_{ug}/t$ as the estimated marginal probabilities of units and groups, respectively.

[1] Agresti (2013) does mention a "dissimilarity index," which is related to the well-known segregation index. However, the "dissimilarity index" mentioned by Agresti was developed by Gini as a measure of the goodness of fit. In the notation of this paper (developed below), this quantity would be $\hat{\Delta} = 1/2 \sum_{u=1}^{U} \sum_{g=1}^{G} |p_{ug} - \hat{t}_{ug}/t|$.

Table 1.1: A contigency table $\mathbf{T}$.

	$g=1$	$g=1$	$\cdots$	$g=G$
$u=1$	t_{11}	t_{12}	$\cdots$	t_{1G}
$u=2$	t_{21}	t_{22}	$\cdots$	t_{2G}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\cdots$
$u=U$	t_{U1}	t_{U2}	$\vdots$	t_{UG}

It is assumed here that the t_{ug} are observed data, from which we compute p_{ug}, $p_{u\cdot}$, and $p_{\cdot g}$, which are estimators of the "true" underlying probability vectors θ_{ug}, $\theta_{u\cdot}$, and $\theta_{\cdot g}$. (These could be parameters of multinomial distributions, for instance.) The counts are assumed to be generated by the probability distributions $f(\theta_{ug})$, $f_u(\theta_{u\cdot})$, and $f_g(\theta_{\cdot g})$, respectively. If the marginal probability distributions $f_u(\theta_{u\cdot})$ and $f_g(\theta_{\cdot g})$ are independent, then the joint distribution $f(\theta_{ug})$ will factor:

$$f(\theta_{ug}) = f_u(\theta_{u\cdot}) f_g(\theta_{\cdot g}) \quad \forall u = 1, \ldots, U, g = 1, \ldots, G.$$

Hence, to assess whether there is a dependency between the two dimensions of the contingency table, we would like to test whether the two marginal distributions are independent. Based on this logic, we can compute the expected cell counts from the data, i.e.

$$\hat{t}_{ug} = t p_{u\cdot} p_{g\cdot}.$$

If the $\hat{t}_{ug}$ are far apart from the t_{ug}, we would reject the independence of the two dimensions. If the $\hat{t}_{ug}$ are close to the t_{ug}—deviating only due to sampling error—we would not reject the hypothesis. This idea gave rise to the Pearson's chi-squared test statistic:

$$X^2(\mathbf{T}) = \sum_{u=1}^{U} \sum_{g=1}^{G} \frac{(t_{ug} - \hat{t}_{ug})^2}{\hat{t}_{ug}}.$$

Under standard assumptions, X^2 is distributed asymptotically as a χ^2 distribution with $(U-1)(G-1)$ degrees of freedom, for reasonably large sample size (Agresti 2013). An alternative is the parametric likelihood ratio test statistic, sometimes also called the G^2 test statistic (Agresti 2013),

$$G^2(\mathbf{T}) = 2\sum_{u=1}^{U}\sum_{g=1}^{G} t_{ug}\log\frac{t_{ug}}{\hat{t}_{ug}},$$

which is also asymptotically distributed as a χ^2 distribution with $(U-1)(G-1)$ degrees of freedom.

Segregation indices are based on a similar idea. If the $\hat{t}_{ug}$ are identical or close to identical to the t_{ug}, then the factors are assumed to be independent and there is no segregation. If the differences between the $\hat{t}_{ug}$ and the t_{ug} are large, then there is a high degree of segregation. The goal of a segregation index is not only to quantify whether there is an association between the two factors *at all*, but also to quantify the strength of the association. One possible segregation index is the Mutual Information Index M, first introduced by Theil and later featured in work by Mora and Ruiz-Castillo (2011). This index is identical to the concept of Mutual Information from information theory as developed by Shannon (1948). Mathematically, the M index is just a rescaled G^2 test statistic:

$$M(\mathbf{T}) := \frac{G^2(\mathbf{T})}{2t} = \sum_{u=1}^{U}\sum_{g=1}^{G} p_{ug}\log\frac{p_{ug}}{\hat{p}_{ug}},$$

where $\hat{p}_{ug} = \hat{t}_{ug}/t$. Clearly, if $p_{ug} = \hat{p}_{ug}$ for all u and g, this index will be minimized at zero, and there is no segregation. Exploiting the connection between the M index and the likelihood ratio test statistic directly suggests a possibility for statistical inference, as $2tM(\mathbf{T})$ will be asymptotically distributed as a χ^2 distribution with $(U-1)(G-1)$ degrees of freedom. Hence, the M index has a direct connection with classical statistics.

What makes the M index less ideal as a measure of segregation is that, unlike most other segregation indices, it is not maximized at 1. When conditioning on the size of the table (i.e., the dimensions $U \times G$), the maximum value the index can reach is $\min\{\log(U), \log(G)\}$. However, when conditioning on the marginal distributions, the maximum value the index can reach is

$E(\mathbf{p}_{\cdot g})$, i.e. the group entropy of the marginal group distribution, $E(\mathbf{p}_{\cdot g}) = -\sum_{g=1}^{G} p_{\cdot g} \log p_{\cdot g}$.[2] The group distribution entropy can then be used to standardize the M index, arriving at the H index:

$$H(\mathbf{T}) := \frac{M(\mathbf{T})}{E(\mathbf{p}_{\cdot g})} = \frac{1}{E(\mathbf{p}_{\cdot g})} \sum_{u=1}^{U} \sum_{g=1}^{G} p_{ug} \log \frac{p_{ug}}{\hat{p}_{ug}}.$$

Because $E(\mathbf{p}_{\cdot g})$ is a function of the data, the direct connection between the index and the G^2 test statistic is broken. I will refer to these two indices, M and H, as entropy-based or information-theoretic segregation indices. In the the next chapter, I discuss the properties of these indices in more detail.

None of the methodological developments in this section imply that the indices M and H (or the test statistics X^2 and G^2, for that matter) can be exclusively applied to the study of segregation. Of course, the chi-squared test was not developed with segregation in mind at all, although it can be reasonably applied in this case. Segregation problems include racial residential, workplace, or school segregation (the unequal distribution of racial groups over geographical units, workplaces, or schools), income segregation (same, but with income groups), and gender occupational segregation. Although these indices are usually only applied to these problems, they can be used for any problem in which it is desired to quantify the dependency between two categorical variables. Therefore, one could call any segregation index also an "index of association." It is used in such a sense in the third chapter, for the study of school-to-work linkages.

This perspective on segregation indices also reveals a similarity to another important literature in sociology, the study of social mobility using discrete categories, for instance, classes. In this case, the dimensions of the contingency table would represent origin and destination classes, for instance. This approach to the analysis of contigency tables would fit models of the form

$$\log \mu_{ug} = \beta + \beta_u^U + \beta_g^G + \beta_{ug}^{UG} \tag{1.1}$$

[2]The conditions for this result to hold are discussed in the following chapter.

to the observed counts of the contingency table, where $\mu_{ug} := t\theta_{ug}$, and the coefficients refer to the unit, group, and interaction parameters, respectively. When testing $\beta_{ug}^{UG} = 0$, this is the same as a test of independence. These models can also be adapted to three-way contingency tables, where the third dimension, for instance, could represent countries. In these models, certain parameters are constrained to be equal across countries, with other parameters varying. The goal, in any case, is to compare the association structure across countries, using parsimonious models that are not confounded by marginal differences between the countries. For instance, assuming a simple 2×2 contingency table, the following relationship for the odds ratio of this table can be derived using eq. 1.1 (Agresti 2013, p. 341):

$$\begin{aligned}
\log \frac{\mu_{11}\mu_{22}}{\mu_{12}\mu_{21}} &= \log \mu_{11} + \log \mu_{22} - \log \mu_{12} - \log \mu_{21} \\
&= \left(\beta + \beta_1^U + \beta_1^G + \beta_{11}^{UG}\right) + \left(\beta + \beta_2^U + \beta_2^G + \beta_{22}^{UG}\right) \\
&\quad - \left(\beta + \beta_1^U + \beta_2^G + \beta_{12}^{UG}\right) - \left(\beta + \beta_2^U + \beta_1^G + \beta_{21}^{UG}\right) \\
&= \beta_{11}^{UG} + \beta_{22}^{UG} - \beta_{12}^{UG} - \beta_{21}^{UG}.
\end{aligned}$$

Hence, the interaction parameters $\{\beta_{ug}^{UG}\}$ solely determine the association, with the marginal coefficients playing no role.

Overall then, the study of segregation and the study of social mobility share similar goals: quantifying the extent of association in contingency tables, ideally by not confounding marginal and structural change. Nonetheless, I am not aware of any work that applies segregation indices to the study of social mobility. In the other direction, there has been one attempt to introduce odds-ratio based segregation measures (which are thus unaffected by changes in the marginal distribution) for the study of occupational gender segregation (Charles 1992; Charles and Grusky 1995; Grusky and Charles 1998). These measures are rarely used, likely because they have several disadvantages compared to the other standard indices. They also have not been discussed in the literature on residential, workplace, or school segregation. In my view, applying segregation indices to the study of social mobility, combined with the marginal-structural decomposition, holds more promise. While not explored in this dissertation, the advantages of such an approach would include a direct measure of the strength of mobility/immobility, and an indicator of how much changes in the marginal distributions affect aggregate changes in social mobility.

1.3 A note on regression analysis in segregation research

A segregation index, by definition, is a "macro" measure: It summarizes the association pattern of a population or subpopulation in a single number. Segregation can only change, however, through change at the "micro" level, i.e., at the individual or household level. The racial segregation of a city can only change through differential fertility or mortality between racial groups, and/or through residential mobility. Linking these micro-level effects to the macro level outcome of interest—segregation—is one of the goals of this dissertation.

Linking the micro and the macro level is inherently difficult with regression models. The usual approach in the literature is to study racial segregation at the level of the metropolitan area. The segregation index of the metropolitan area is then regressed on characteristics, either in the cross-section or in a panel dataset, where the characteristics are time-varying and metropolitan area fixed effects are added to the model. Thus, these models attempt to find factors that are associated either with absolute levels of segregation, or with changes in segregation. There are at least three major problems with this approach.

First, what is common to the regression-model approach is that the explanatory power can only come from changes or characteristics at the level of the metropolitan area. These "macro explanations" are therefore disconnected from the more localized theories that are usually cited to explain segregation in the U.S. For instance, differences in socioeconomic status between racial groups are large in the U.S., and several studies find that metropolitan areas in which the income differences between racial groups are smaller tend to have lower segregation as well (Logan et al. 2004; Reardon et al. 2009). However, these differences cannot fully explain patterns of residential segregation, and this is especially true for Blacks, who remain more segregated than their economic status would predict (Galster and Sharkey 2017). This suggests that

Black families experience ongoing economic disadvantage, which hinders their moving out of disadvantaged neighborhoods. While such arguments are intriguing, they are hard to test at the level of the metropolitan area, and it is therefore hard to disentangle the importance of different explanations.

Second, by regressing an aggregate segregation index on other aggregate measures, regression models at the level of the metropolitan area are possibly subject to the ecological fallacy. For instance, assume that wealthy individuals are more likely to segregate, and this is tested by including a predictor measuring average wealth of the metro area in the regression model. The problem with this approach is that it can still be the case that wealthy metro areas are less segregated. The relationship depends on the distribution of wealth, and if wealthy metro areas have higher income inequality, they will also be less segregated. To actually assess the impact of wealthy individuals on segregation, individual-level mobility behavior would need to be observed.

Third, the regression model may not be successful in explaining the processes that we would like to understand. Assume for instance, that segregation declined everywhere at a similar rate, and that the decline is mostly explained by Black suburbanization. As will be shown in chapter 5, this scenario is not an entirely uncharacteristic representation of the actual situation. The problem is that there is not much variability in the treatment (Black suburbanization), such that the regression model will not be able to identify it as an important factor. As Bhrolcháin and Dyson (2007) put it, in such a situation the "true explanandum is the uniformity of decline rather than its variability" (p. 5).

In this dissertation, I therefore take a different approach to understanding changes in segregation, which involves decomposition analysis based on simulations (discussed in detail in chapter 4). These decompositions establish the localized link between micro-level neighborhood change and macro-level segregation outcomes more directly, and are hence able to identify the important population movements that influence changes in segregation.

Chapter 2: A Method for Studying Differences in Segregation Across Time and Space

2.1 Introduction

Studies of segregation are concerned with a variety of substantive problems.[1] Social scientists are interested in residential racial segregation, in the racial or class-based segregation of schools and workplaces, or the gender segregation of occupations. More generally, any study of the association between two categorical variables can be regarded as a segregation problem. Segregation is usually studied by applying a segregation index to a contingency table, which provides a one-number summary of the association between, for instance, gender and occupations.

Often, the interest in the study of a segregation problem lies not only in describing segregation at one point in time or in one place, but in comparing levels of segregation over time, across countries or cities, or between population groups. For instance, in the school segregation literature, there is a debate about the resegregation of schools along racial lines (Reardon and Owens 2014). The workplace segregation literature documented a decrease in within-workplace racial segregation levels, but a decrease in between-workplace segregation (Ferguson and Koning 2018). The gender-occupational literature is interested not only in comparing segregation over time within a single country, but also across regional or national economies (e.g., Charles and Grusky 2004). When comparing across time, the message of segregation studies is often that segregation has either increased or decreased, but the deeper causes for these differences often remain unclear. The contribution of this paper is to provide a general and

[1] A slightly adapted version of this chapter was published as: Elbers, Benjamin. 2021. "A Method for Studying Differences in Segregation Across Time and Space." *Sociological Methods & Research*. `doi: 10.1177/0049124121986204`.

practical method for the study of change or difference in segregation (hereafter abbreviated as "change" when this doesn't create confusion). The method developed here brings practical advantages to many segregation problems, and proposes a solution to the long-standing problem of margin-dependency:

1. Among the practical advantages, the method allows an analysis of where differences in segregation originate. For instance, we might ask whether declining occupational gender segregation arises from manual or professional occupations, or whether the declines in segregation are associated with changes in the educational composition of certain occupations. In school or residential segregation, it would be of interest to know which schools or neighborhoods contribute most to changes in segregation. This is especially relevant when considering different types of schools, such as charter or private schools. Relatedly, one may study the association between gentrification and segregation at the neighborhood level.

2. The method also allows for a change in the number of units under study. This problem arises naturally in the study of school segregation: When comparing school segregation across two points in time, some schools will have closed and new ones will have opened. The problem may also occur with occupational segregation: Over time, some occupations will become obsolete and vanish, while new occupations appear. The method developed here allows the researcher to quantify the effect of these "appearing" and "disappearing" units on the total change in segregation. While this seems to be a natural question, it has only received scant attention in the segregation literature. An exception is Ferguson and Koning (2018), who studied the effect of firm turnover on workplace segregation.

3. Lastly, the method provides a solution to the problem of margin-dependency. Taking again occupational gender segregation as an example, it is intuitively clear that some of the declines in gender segregation of recent decades may be due to compositional changes. Deindustrialization has led to declines in factory jobs and a decrease in the share

of manual and routine occupations, which have often been almost entirely male (Weeden 2004). If these occupations are still as segregated as they were before, and only their relative share has declined, this will register as a decrease in (most) segregation indices. Thus, it would be desirable to distinguish between these changes, which are referred to as marginal changes (because the change is reflected in the marginal row or column sums of the contingency table), from changes in "pure segregation." A major part of this article will elaborate on this distinction and on the exact meaning of "pure segregation."

The methodological literature on segregation indices has engaged mostly with point (3), the margin-dependency of segregation indices, while the useful innovations described in points (1) and (2) have received almost no attention. The method described in this paper proposes a solution to the margin-dependency problem that can be summarized as follows: Margin-dependency is desirable in the cross-section to characterize the "average" level of segregation an individual experiences, but is problematic when comparing levels of segregation across time or space. The solution, as first proposed by Karmel and Maclachlan (1988), is to decompose the difference into terms that distinguish changes that are introduced because of changes in the marginal distributions alone from changes in "pure segregation." The latter will be called "structural change" throughout the paper. Combining this idea with the desired properties of (1) and (2), we arrive at a five-term decomposition:

$$\begin{aligned} S(t_2) - S(t_1) = {} & \Delta_{\text{appearing}} + \Delta_{\text{disappearing}} \\ & + \Delta_{\text{marginal-units}} + \Delta_{\text{marginal-groups}} \\ & + \sum_{u \in t_1 \cap t_2} \Delta_{u,\text{structural}} \end{aligned}$$

where $S(\cdot)$ refers to the value of the segregation index at different points in time. The equation then says that we decompose the difference in segregation between two time points (or across population groups, places) into two terms that account for the appearance and disappearance of units under study (think school openings and school closures), two terms that account for compositional changes (the marginal distributions), both in terms of units (say, schools) and groups (say, racial groups). The last term is a summation that extends over those units that are present at both time points, and describes the change in structural (or "pure") segregation that arises from each unit.

Thus, the decomposition opens up new avenues of research for scholars working on segregation problems. Its primary advantage is that it allows for a much more precise statement about the nature of change: We can pinpoint whether the segregation change is due to a change in the population of units, due to marginal change, or due to structural change. We can further drill down to study whether the structural change is concentrated in a certain set of units that are of special interest (say, charter schools). It should also be noted that the total change in segregation could be zero, but that some of the components are non-zero. In this case, some positive components would be offset by negative components. The decomposition of change could thus reveal previously obscured patterns, such as an increase in "pure segregation" that is offset by declines due to marginal changes. (Such offsetting patterns are explored in the examples at the end of the paper.)

The remainder of the paper is organized as follows: In the next section, the issue of margin-dependency and the possible solutions that have been presented in the literature are discussed. None of these solutions are deemed satisfactory. It is then argued that the only index that can fully achieve the desired five-term decomposition is the M index. This index, extensively discussed by Mora and Ruiz-Castillo (Mora and Ruiz-Castillo 2003, 2009, 2011), is not as widely used as the closely-related H index, but has many desirable properties. Next, the decomposition procedure is introduced. Lastly, the practical advantages of the method are shown through two examples: changing occupational gender segregation in the U.S. and changing residential segregation in Brooklyn, New York.

2.2 The problems and benefits of margin dependency

To make the following more concrete, consider U organizational units, such as schools or occupations, and a number of population groups, G, such as racial groups or genders. For an occupational segregation problem, the number of workers in each occupation-gender combination can be cross-classified in a $U \times G$ contingency table. A segregation index $S(\cdot)$ is a function that summarizes the $U \times G$ contingency table to a single number. Without loss of generality, occupational gender segregation will be used as an example for the remainder of the paper.

Margin-dependency refers to the property of some segregation indices that proportional changes in the marginal distributions of the contingency table lead to a change in the index value. To illustrate, consider a simplified economy with three occupations and two genders. At time point 1, there are 55 men and 45 women, distributed across occupations in a way that the first occupation is integrated, while the other two are rather segregated. This matrix is shown at the left-hand sides of the arrows, with men in the first and women in the second column:

$$t_1 : \begin{bmatrix} 25 & 25 \\ 28 & 2 \\ 2 & 18 \end{bmatrix} \qquad \rightarrow \qquad t_2 : \begin{bmatrix} 20 & 20 \\ 28 & 2 \\ 4 & 36 \end{bmatrix}$$

$$t_1 : \begin{bmatrix} 25 & 25 \\ 28 & 2 \\ 2 & 18 \end{bmatrix} \qquad \rightarrow \qquad t_2^* : \begin{bmatrix} 25 & 50 \\ 28 & 4 \\ 2 & 36 \end{bmatrix}$$

Consider then two alternative scenarios. In the first scenario (top), the size of the first, integrated occupation decreases by 20%, and the third occupation (which is very segregated) doubles. Note that it is not possible under these transformations to keep the gender proportion constant without changing the internal proportion of the remaining occupation. In the second scenario (bottom), the size of the female labor force doubles, with the numbers for men unchanged. An index that changes its value under the first transformation is called unit-margin-dependent, while an index that changes its value under the second transformation is called group-margin-dependent. A margin-free index, by definition, does not change under either of these processes. An overview of prominent indices is displayed in Table 2.1.

The entropy-based, information-theoretic indices M and H are margin-dependent for both groups and units. This is also true for the variance ratio index V (also known as separation or eta-squared index). Other indices, such as the index of dissimilarity D, are only margin-

Table 2.1: Margin-dependency of different indices

		Unit (e.g., Occupation)	
		margin dependent	**margin free**
Group	**margin dependent**	*M*, *H*, *V*	*SSD*
(e.g., Gender)	**margin free**	*D*	*A*

Note: Adapted from Charles and Grusky (1995, p. 934)

dependent in terms of the unit distribution. The size-standardized index of dissimilarity *SSD* is group-margin-dependent only,[2] and only the log-linear index *A* is margin-free in both dimensions.

The reader might be surprised to find the *H* index among the group-dependent indices. The margin dependency of the *H* index is often not explicitly considered in empirical studies, although this fact is known at least since James and Taeuber (1985). For instance, An and Gamoran (2009, p. 20) write that they "use a measure [the *H*] that is insensitive to changes in the U.S. school population, thereby concentrating solely on racial imbalance." This, however, is not entirely true. While the *H* index involves a term that partly accounts for changes in group marginals, the standardization is not complete (for a formal proof, see Mora and Ruiz-Castillo 2011). We thus emphasize here that the *H* is margin-dependent in both directions.

Often, margin-dependency is considered problematic, and the segregation literature has devoted considerable effort to solving this problem. The problem stems from the assumption that marginal changes often reflect processes that are thought to be unrelated to the deeper, structural causes of segregation. For instance, deindustrialization (changing occupation marginals) or a rising share of female employment (changing group marginals) should only lead to changes in segregation if the structure of segregation changed. If the changes are in the margins only, arguments about "rising" or "decreasing" segregation may not be warranted.

The major solutions to the problem are discussed in turn:

[2] The unit-margin-dependency of the *D* has long been recognized. This led to the development of the size-standardized index of dissimilarity *SSD*, which, while not margin-dependent on the unit distribution, reintroduces a dependency on the group distribution.

1. A series of papers by Charles and Grusky (Charles 1992; Charles and Grusky 1995; Grusky and Charles 1998) introduced the A index. The A index is based on the insight that a measure of segregation that is invariant to row or column transformations needs to be based on odds ratios. For instance, two local odds ratios are sufficient to describe the association structure of a 3×2 table (as in the example above). If we let n_{ij} denote the number of workers in the ith row and jth column, the two odds ratios are $\theta_{1,1} = \frac{n_{1,1} n_{2,2}}{n_{1,2} n_{2,1}}$ and $\theta_{2,1} = \frac{n_{2,1} n_{3,2}}{n_{2,2} n_{3,1}}$ (Agresti 2013, p. 54). It is easy to verify that these odds ratios are identical for all three matrices t_1, t_2 and t_2^*, which is to say that the association structure between occupations and gender does not change from t_1 to t_2 or from t_1 to t_2^*. This is the same argument that is made in favor of log-linear modeling in the study of social mobility.

   Essentially, the A index calculates the odds ratio of male and female employment within each occupation, and is then summarized by weighting all occupation-specific ratios equally. The resulting index measures only the level of association as captured by the odds ratios, and is not influenced by changes in the marginal distribution of either occupations or genders. Note that the index achieves its unit-margin-independence by simply weighting all occupations equally. The index is thus more a characterization of the segregation of the average occupation, and not a measure of average segregation at the individual level. Especially if the sizes of occupations differ greatly, the index is problematic. (See also the exchange between Watts (1998) and Grusky and Charles (1998).) It thus seems even less applicable when school or residential segregation is studied.

   Another way to phrase this problem is that the A index conflicts with the criterion of organizational equivalence. Organizational equivalence implies that when two occupations with the same level of segregation are combined, segregation should be unchanged (James and Taeuber 1985). This criterion is not fulfilled when occupations are weighted equally and the segregation level of the other, uncombined occupations differ from the two occupations that are combined. This shows that the discussion about the merits of margin-free versus margin-dependent indices cannot be resolved, because the two indices pursue goals that are not compatible.

2. Karmel and Maclachlan (1988) propose a decomposition that is very similar to the one developed in this paper. Their approach is based on creating counterfactual contingency tables that account only for the effects of marginal and structural changes, respectively. This is done using iterative proportional fitting (IPF), which will be explained below. The counterfactual tables can then be used to disentangle marginal from structural changes. A downside of their approach is that the decomposition contains an interaction effect between the two marginal dimensions, which is hard to interpret. They also do not address the problem of appearing and disappearing units. The largest disadvantage of their method is the choice of index, which they call I_p, and which is not decomposable in terms of units or groups.

3. Mora and Ruiz-Castillo (2009) presented two formulas that supposedly quantify structural and compositional change between two M indices. With a slightly adapted notation, the difference between two M indices, defined by the matrices t_1 and t_2, is decomposed as follows:

$$\begin{aligned} M(t_2) - M(t_1) &= \Delta\mathrm{N}(\Pi^u) + \Delta\mathrm{G}^u + \Delta\mathrm{U}(\Pi^u) \\ &= \Delta\mathrm{N}(\Pi_g) + \Delta\mathrm{U}_g + \Delta\mathrm{G}(\Pi_g), \end{aligned} \tag{2.1}$$

where the ΔU and ΔG capture changes in the marginals of unit and group proportions, respectively, and ΔN captures "composition-invariant" changes, which, importantly, are not the same as structural changes defined through the change in odds ratios. As the authors themselves write, the interpretation of these terms hinges on crucial assumptions which are rarely met in practice (Mora and Ruiz-Castillo 2009; Watts 2015, p. 47f.). For reasons of brevity, these problems are not explicated fully here. Instead, an especially problematic aspect of these decompositions is highlighted: There are two possible answers for each of the three components, and these will in most cases provide conflicting interpretations. The decompositions on the first and the second line will only in exceptional circumstances give the same numerical results. This is easily seen by applying equation (2.1) to the difference between t_1 and t_2 from the example above:

$$\begin{aligned} M(t_2) - M(t_1) &= 0 + 0.00376 + 0.0267 \\ &= -0.0209 + 0.0479 + 0.00346 = .03 \end{aligned}$$

The first decomposition implies that structural change is zero, and further suggests that the marginal change in the occupational distribution is largely responsible for the increase in segregation, which aligns with our expectations. However, the second line gives a contradictory answer, implying that structural segregation decreased (−0.0209). Furthermore, the size of the marginal components is not the same in the two decompositions. Even if the assumptions that underlie these decompositions were justified in practice (which is questionable), the fact that the two decompositions give two possibly contradictory answers is unsatisfactory and poses practical problems of interpretation. The issue here that these decompositions are not based on the notion that only the odds ratios are invariant under row and column transformations.

Lastly, their decomposition also does not address the problem of appearing and disappearing units, which means that only the common subset of units can be decomposed.

The method developed in this paper is based on the idea that margin-dependency (especially in terms of units) is a desirable property in the cross-section. Consistent with the idea that we want to measure average segregation at the individual level, it is reasonable to argue that a segregation index should be higher when more people work in segregated occupations. If occupations are weighted equally, this is not the case. At the same time, we would also like to distinguish changes that are purely due to composition (marginal changes) from changes in pure segregation (structural changes). To illustrate this point, consider that two processes occur at the same time: the occupations that are more segregated grow at the expense of less-segregated occupations, while at the same time segregation *within* each occupation declines. The overall change in segregation will be positive if the first process leads to a greater change than the second process. If attention is only paid to the total difference, the conclusion will be that segregation has become "worse" (which is a warranted statement, at least for the average worker). However, the statement is also imprecise, because the segregation of each individual occupation has in fact decreased. The decomposition of change into the two components thus allows the researcher to pinpoint more clearly the sources of segregation change. Importantly, the prevalence and direction of the two trends may call for different policy responses.

Thus, the paper advocates for an approach that uses a margin-dependent index in the cross-section, which is then decomposed when we compare over time or across places. The proposed solution combines and expands the approaches (1)-(3) discussed above. Charles and Grusky provide the key insight that any structural changes are reflected in the odds ratios, and that these are the only measures of association that are invariant under marginal transformations. Karmel and Maclachlan use iterative proportional fitting to arrive at counterfactual tables. Lastly, Mora and Ruiz-Castillo's contributions highlighted the advantages of the entropy-based index M, which will be adopted below.

2.3 The choice of index

Recently, the H has become increasingly popular for the study of racial segregation, which is most likely due to two distinct advantages. First, the H allows for attractive decompositions. Second, the H allows for a natural treatment of the multigroup case, which has become increasingly important for the study of racial segregation in the U.S., and is a natural requirement in other segregation problems. In their comprehensive overview of multigroup segregation indices, Reardon and Firebaugh (2002) conclude "that the information theory index H is the most conceptually and mathematically satisfactory index" (p. 33).

In a recent series of papers, Mora and Ruiz-Castillo (2003, 2009, 2011) pointed to an alternative, but closely related index, which they called the *Mutual Information Index (M)*. Both the

M and H were introduced by Theil (Theil 1967, 1971, 1972; Theil and Finizza 1971). Mora and Ruiz-Castillo, as well as Frankel and Volij (2011), outlined some of the advantages of the M over the H. Importantly, Mora and Ruiz-Castillo (2011) showed that the decomposition of an H index into between- and within-group terms (for instance, white/non-white) may be ambiguous, and they thus recommend the adoption of the M if such decompositions are desired.[3]

To define H and M, assume that we observe the gender composition of U occupations. Define t_{ug} as the number of workers with gender g in occupation u, and the total number of workers as t. From this contingency table, define $p_{u\cdot} = \sum_{g=1}^{G} t_{ug}/t$ and $p_{\cdot g} = \sum_{u=1}^{U} t_{ug}/t$ as the marginal probabilities of occupations and gender, respectively. The joint probability of being in occupation u and gender g is $p_{ug} = t_{ug}/t$. We also write $p_{g|u} = p_{ug}/p_{u\cdot}$ as the conditional probability of having gender g given occupation u (and $p_{u|g}$ likewise).

The M index quantifies how strongly each occupation's gender distribution deviates from the overall (or expected) gender distribution. This yields a "local" segregation score for each occupation, called L_u. The occupation scores are then weighted by the size of the occupation, $p_{u\cdot}$. To measure the deviation, the logarithm of the ratio between conditional and marginal probabilities is used. As Theil (1972) has shown, the logarithm allows for the attractive decomposition properties. Thus,

$$M = \sum_u p_{u\cdot} L_u = \sum_u p_{u\cdot} \left(\sum_g p_{g|u} \log \frac{p_{g|u}}{p_{\cdot g}} \right).$$

Because the M is symmetric, it can also defined by summing proportion-weighted scores for each gender, i.e.

$$M = \sum_g p_{\cdot g} L_g = \sum_g p_{\cdot g} \left(\sum_u p_{u|g} \log \frac{p_{u|g}}{p_{u\cdot}} \right).$$

The simple expressions for the M, i.e.

[3] Mora and Ruiz-Castillo (2011, p. 161) identify a small number of papers that prefer the M index over the H. Beyond those, DiPrete et al. (2017) and Forster and Bol (2018) have used the M index in the context of school-to-work linkages.

$$M = \sum_u p_{u\cdot} L_u = \sum_g p_{\cdot g} L_g$$

show that the M is symmetric (i.e., the meaning of groups and units can be exchanged), and that the M is margin-dependent in both directions. From the standpoint of decomposing changes in segregation, this is an attractive property.

The M can also be motivated from an information-theoretic perspective, which is helpful to understand its basic properties. First, define the entropy $E(\cdot)$ of a distribution as

$$E(\mathbf{p}) = -\sum_i p_i \log p_i,$$

where $\mathbf{p}$ is a vector of probabilities that sums to 1. Entropy is a non-negative measure of *expected information* or *uncertainty* (Theil 1972). Consider two events that occur with probabilities .99 and .01. The expected information of the next observation from this distribution is close to zero, i.e. $E([.99, .01]) = .06$, as we were virtually certain that the first event would occur. However, for two events that will occur with a probability of 1/2 each, the expected information is large, i.e. $E([.5, .5]) = \log 2 \approx .69$. The entropy is maximized at $\log n$ when the probability of each event is $1/n$, where n is the number of events. Intuitively, the entropy is minimized at zero when it is certain which event will occur.

To define M from this perspective, we ask: How much more information does the overall distribution provide compared with the gender distribution of a specific occupation? Formally, this is the difference in entropies at the occupation level, weighted by the occupation's proportion:

$$M = \sum_u p_{u\cdot} [E(\mathbf{p}_{\cdot g}) - E(\mathbf{p}_{g|u})],$$

where $\mathbf{p}_*$ refers to the relevant vector of probabilities. Due to the symmetry of the M, this expression can also be formulated from the gender perspective:

$$M = \sum_g p_{\cdot g} [E(\mathbf{p}_{u\cdot}) - E(\mathbf{p}_{u|g})].$$

It follows that M is minimized at zero when the gender distribution of each occupation is identical to the overall gender distribution. M is maximized at $\min(\{\log U, \log G\})$. To see this, note that Equation (2.3) is maximized when the entropy $E(\mathbf{p}_{\cdot g})$ is maximized, and the entropy $E(\mathbf{p}_{g|u})$ for each occupation is minimized. This is the case when each gender has the same overall proportion, and when each occupation is either completely male or completely female.

It may seem odd that a segregation index can only be maximized when all groups are the same size, but it is in line with information-theoretic principles. This point will become clearer with an example. Consider two labor markets A and B with 200 workers each, and only three occupations. The labor markets differ in their gender distributions. Labor market A has 100 women and 100 men, while B has 20 women and 180 men. The workers are distributed as follows, with the occupations indexed by the rows of the matrix:

$$
\begin{array}{cc}
 & \begin{array}{cc} \text{women} & \text{men} \end{array} \\
\text{A:} & \begin{bmatrix} 100 & 0 \\ 0 & 50 \\ 0 & 50 \end{bmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \begin{array}{cc} \text{women} & \text{men} \end{array} \\
\text{B:} & \begin{bmatrix} 20 & 0 \\ 0 & 90 \\ 0 & 90 \end{bmatrix}
\end{array}
$$

In both labor markets, all three occupations are completely segregated in the sense that there is no mixing within occupations. For these matrices, $M(A) = .69$ and $M(B) = .33$. The M index thus finds that segregation in A is twice as high as in B. This suggests to standardize the M index by the gender entropy, which gives the H index:

$$H = \frac{M}{E(\mathbf{p}_{\cdot g})}.$$

For the two cities, it follows that $H(A) = H(B) = 1$. The H is attractive because it is standardized between zero and one,[4] which facilitates comparisons between two cities with differing

[4]Three caveats apply: First, the standardization is only limited to the range from zero to one when $U \geq G$, which is the case in most segregation problems. Alternatively, Mora and Ruiz-Castillo (2011) also define the H^* index. This index is defined by standardizing the M by the unit distribution entropy, i.e. $H^* = \frac{M}{E(\mathbf{p}_{u\cdot})}$. This index is maximized when $E(\mathbf{p}_{u|g}) = 0$ for all groups, which is only true when all members of each group are concentrated at exactly *one* unit. This, of course, is not possible with two groups and more than two units. The H^* index is thus only appropriate when $G \geq U$, which for practical segregation problems is usually not the case. Second, the maximal value can only be reached if there are more subjects than there are units. As Carrington and Troske (1998, p. 239) write, "in a sample with 10 black workers and 20 firms, for example, evenness is unobtainable because

gender distributions. Nonetheless, there is an argument to be made for the M index. While the H index sees the amount of segregation as equal between the two cities, the M takes into account that it is much "harder" in A to achieve complete segregation than it is in B. Given that in B 90% of the workers are men, it is less surprising to find an all-men occupation in B than it is in A.

2.4 The decomposition of change

2.4.1 Generating counterfactuals through IPF

Instead of attempting the margin-free measurement of segregation at each point in time, the approach outlined here follows the idea that *changes* in segregation indices can be decomposed into marginal and structural changes (Mora and Ruiz-Castillo 2009; Watts 1998, 2015). This method was proposed by Theil himself (Theil 1972, p. 131ff.), and was extended by Karmel and Maclachlan (1988) in the context of occupational gender segregation. Karmel and Maclachlan used another segregation index, but the approach is applicable whenever a margin-free comparison of two contingency tables is desired.

The basic idea is to adjust the contingency table from time point t_1 forward so that only marginal changes between the two time points are taken into account. Consider a labor market with men and women distributed across three occupations. We observe the labor market at two points in time. Between these two time points, the number of men has grown and the number of women has declined. At the same time occupations have changed in size, with especially strong declines in the third occupation. The question is: if there are changes in segregation, how much of these changes can be attributed to changes in the distribution of gender and occupation marginals alone, and how much of the change can be attributed to changes in the odds ratios?

it is impossible for each firm to get half a black worker." In many practical segregation problems, this is usually not a problem as there are more subjects than units. Third, the standardization only works when the size of the smallest group is larger than the smallest unit. For instance, consider a labor market of 200 women and 700 men, distributed across three occupations of size 300. Even if the occupations are maximally segregated (i.e. two all-men occupations, and one occupation with 200 women and 100 men), the indices reach their maxima at $M = .32$ and $H = .6$. Whether such marginal constraints matter in practice depends on the concrete application.

At the two time points, the workers are distributed across occupations as follows:

$$
t_1 : \begin{array}{c} \begin{array}{cc} \text{women} & \text{men} \end{array} \\ \begin{bmatrix} 20 & 100 \\ 180 & 50 \\ 600 & 50 \end{bmatrix} \end{array}
\qquad
t_2 : \begin{array}{c} \begin{array}{cc} \text{women} & \text{men} \end{array} \\ \begin{bmatrix} 10 & 170 \\ 80 & 60 \\ 240 & 40 \end{bmatrix} \end{array}
$$

Both the M and the H register large changes in segregation: the M increases by over 80% between t_1 and t_2, while the H increases by 33%. To identify how much of this change is due to marginal changes, the matrix at t_1 is transformed to have the same margins as t_2, while leaving the association structure (i.e., the odds ratios) intact. This can be achieved using Iterative Proportional Fitting (IPF): First, the cells of t_1 are scaled to achieve the overall *gender* marginal distribution of t_2. The adjusted cell counts are then scaled to achieve the marginal *occupation* distribution of t_2. This process is repeated until the margins of the adjusted table are within a small percentage of t_2. The first steps of the procedure are shown here:

$$
\begin{bmatrix} 20 & 100 \\ 180 & 50 \\ 600 & 50 \end{bmatrix} \Longrightarrow \begin{bmatrix} 20 \times 330/800 & 100 \times 270/200 \\ 180 \times 330/800 & 50 \times 270/200 \\ 600 \times 330/800 & 50 \times 270/200 \end{bmatrix}
$$

$$
= \begin{bmatrix} 8.3 & 135 \\ 74.3 & 67.5 \\ 248 & 67.5 \end{bmatrix} \Longrightarrow \begin{bmatrix} 8.3 \times 180/144.4 & 135 \times 180/144.3 \\ 74.3 \times 140/141.8 & 67.5 \times 140/141.8 \\ 248 \times 280/315.5 & 67.5 \times 280/315.5 \end{bmatrix}
$$

$$
= \begin{bmatrix} 10.3 & 168.4 \\ 73.4 & 66.6 \\ 220.1 & 59.9 \end{bmatrix} \Longrightarrow \begin{bmatrix} 10.3 \times 330/303.8 & 168.4 \times 270/294.9 \\ 73.4 \times 330/303.8 & 66.6 \times 270/294.9 \\ 220.1 \times 330/303.8 & 59.9 \times 270/294.9 \end{bmatrix}
$$

$$
= \begin{bmatrix} 11.2 & 154.2 \\ 79.7 & 61 \\ 239.1 & 54.8 \end{bmatrix} \Longrightarrow \begin{bmatrix} 11.2 \times 180/165.4 & 154.2 \times 180/165.4 \\ 79.7 \times 140/140.8 & 61 \times 140/140.8 \\ 239.1 \times 280/293.9 & 54.8 \times 280/293.9 \end{bmatrix}
$$

$$
= \begin{bmatrix} 12.2 & 167.8 \\ 79.2 & 60.7 \\ 227.8 & 52.2 \end{bmatrix} \Longrightarrow \text{... (10 steps omitted)}
$$

$$
= \begin{bmatrix} 13.7 & 166 \\ 83.5 & 56.5 \\ 233 & 47.3 \end{bmatrix} = t_1'
$$

The transformations at rows one and three adjust the gender marginals; while the transformations at rows two and four adjust the occupation marginals. It is unimportant whether the procedure starts with the group or the unit marginals; it will always converge (for details on IPF, see Agresti 2013; Deming and Stephan 1940).[5] After four steps, both margins are already within 3-4% of the desired marginals. After 14 steps, the procedure yields the matrix shown in the last row, where the marginals are within 0.1% of the desired marginals. The resulting matrix t_1' is a counterfactual version of the t_1 matrix, where only the marginals changed in the direction empirically observed in t_2, but the odds ratios are the same as in t_1. This allows a decomposition of overall change in segregation levels as follows:

$$\begin{aligned} M(t_2) - M(t_1) &= \overbrace{M(t_2) - M(t_1')}^{\text{structural}} + \overbrace{M(t_1') - M(t_1)}^{\text{marginal}} \\ &= (.273 - .238) + (.238 - .150) \\ &= .035 + .088 = .123 \end{aligned} \tag{2.2}$$

The "marginal" component quantifies how much we would expect segregation to change given that the marginals changed towards those of t_2. The "structural" component quantifies any additional amount of segregation that is unexplained by marginal changes. To understand the behavior of the decomposition, it is useful to consider the two extreme cases of "structural change only" and "marginal change only". Considering t_1, it is possible to construct an alternative matrix that redistributes the workers across occupations in such a way that the marginals will stay the same (e.g., by distributing 50 workers from occupation 1 to the other two occupations, and moving the same number of women to occupation 1.) A decomposition of these two matrices will find that marginal change is zero, because the IPF procedure converges immediately without changing any cell counts. Thus, the marginal term of Equation (2.2) would compare identical matrices, and the difference would be zero—as desired. Similarly, it is also possible to construct a matrix where simply the number of, say, women doubled. In this case, the IPF procedure scales the margins in exactly this way, which means that the structural term of Equation (2.2) compares identical matrices, and we get again the desired result.

[5] The IPF procedure requires positive counts in each cell, which in practice may not always be the case. The canonical solution here is to replace zero counts with a very small number, e.g., 0.0001.

One criticism that can be leveled against this decomposition is that the choice of t_1 as the baseline is somewhat arbitrary, especially if the matrices are not compared over time, but across space or, say, across birth cohorts. The results are similar, but not identical when we instead choose t_2 as the baseline and apply the IPF procedure to this matrix:

$$\begin{aligned} M(t_2) - M(t_1) &= \overbrace{M(t_2') - M(t_1)}^{\text{structural}} + \overbrace{M(t_2) - M(t_2')}^{\text{marginal}} \\ &= .026 + .097 = .123 \end{aligned}$$

In decomposition analysis, this is known as the path-dependency problem (Fortin et al. 2011; Kitagawa 1955), where the results of the decomposition are dependent on the order in which elements are eliminated. As proposed by Shorrocks (2013), the solution to this problem is the Shapley decomposition, which considers all possible ways in which an element can be eliminated. In this case, the decomposition results in a simple averaging of the two scenarios (Deutsch et al. 2009):

$$\begin{aligned} M(t_2) - M(t_1) &= \overbrace{\frac{1}{2}(M(t_2) - M(t_2')) + \frac{1}{2}(M(t_1') - M(t_1)}^{\Delta_{\text{marginal}}} \\ &\quad + \underbrace{\frac{1}{2}(M(t_2) - M(t_1')) + \frac{1}{2}(M(t_2') - M(t_1))}_{\Delta_{\text{structural}}} \end{aligned} \tag{2.3}$$

For the example, this is

$$M(t_2) - M(t_1) = \underbrace{\frac{1}{2}(.097 + .088)}_{\Delta_{\text{marginal}}} + \underbrace{\frac{1}{2}(.035 + .026)}_{\Delta_{\text{structural}}} = .092 + .031 = .123.$$

From this decomposition, we conclude that marginal changes are responsible for about three quarters of the overall change in the M, while structural changes account for only a quarter of the increase. Compared to segregation indices that focus on structure only (i.e., odds ratios), the procedure introduced here quantifies the effects of both marginal and structural changes. It will be argued below that marginal changes are often an important part of segregation processes, and that it is therefore not always desirable to “purge” the influence of the marginal distributions.

This aggregate view of segregation differences can be further decomposed. The key property that is exploited here is that in the marginal component, the odds ratios are the same, and that in the structural component, the marginal distributions of units and groups are the same.

2.4.2 Decomposing marginal changes further

The marginal change can be further subdivided into two components: one component quantifies the contribution of changing unit marginals and one quantifies the contribution of changing group marginals. Karmel and Maclachlan proposed a simpler decomposition that includes an interaction term, but the Shapley decomposition can be used to quantify the contributions of either margins without an interaction term. A full proof of this strategy is provided by Deutsch et al. (2009), and I will present here the intuitive understanding of this decomposition. Again, consider all the ways in which either marginal component can be eliminated. For this, consider all possible combinations between unit marginals, group marginals, and odds ratios from both t_1 and t_2. As a shorthand notation, write $M(U;G;O)$ to identify the M that is calculated based on the unit (row) marginals from U, the group (column) marginals from G, and the odds ratios from O. For instance, $M(t_1) = M(t_1;t_1;t_1)$ and $M(t_1') = M(t_2;t_2;t_1)$. Given all possible combinations, there are eight unique matrices, including the two unaltered ones. This decomposition thus requires six distinct IPF procedures. For instance, to arrive at $M(t_1;t_2;t_1)$, the matrix t_1 has to be adjusted towards the column marginals of t_2 while retaining its original t_1 row marginals. The decomposition then relies on averaging all possible elimination strategies. To quantify the effect of marginal change in the rows, there are four possible elimination strategies:

$$\begin{aligned}\Delta_{\text{marginal-units}} = &\frac{1}{4}(M(t_2;t_1;t_1) - M(t_1;t_1;t_1)) + \frac{1}{4}(M(t_2;t_2;t_1) - M(t_1;t_2;t_1)) \\ &+ \frac{1}{4}(M(t_2;t_2;t_2) - M(t_1;t_2;t_2)) + \frac{1}{4}(M(t_2;t_1;t_2) - M(t_1;t_1;t_2))\end{aligned} \tag{2.4}$$

Note that within each subtraction, only the row margins are changed, with the other two factors held constant. Similarly, for the columns:

$$\begin{aligned}\Delta_{\text{marginal-groups}} = &\frac{1}{4}(M(t_1;t_2;t_1) - M(t_1;t_1;t_1)) + \frac{1}{4}(M(t_2;t_2;t_1) - M(t_2;t_1;t_1)) \\ &+ \frac{1}{4}(M(t_2;t_2;t_2) - M(t_2;t_1;t_2)) + \frac{1}{4}(M(t_1;t_2;t_2) - M(t_1;t_1;t_2))\end{aligned} \tag{2.5}$$

Simple algebra shows that $\Delta_{\text{marginal-units}} + \Delta_{\text{marginal-groups}} = \Delta_{\text{marginal}}$. Applying this decomposition to the example above, we get:

$$\begin{aligned}\Delta_{\text{marginal}} &= \Delta_{\text{units}} + \Delta_{\text{groups}} \\ &= .082 + .01 = .092\end{aligned}$$

Among the changes in the marginals, the shift in the unit marginals was much more important for the increase in segregation than the shifting gender distribution, despite the large changes.

2.4.3 Decomposing structural changes further

Usually, structural change is of greater interest than marginal change. The term for the structural component admits two straightforward decompositions based on local segregation scores. These decompositions were not exploited by Karmel and Maclachlan (1988) or others, because their index did not admit disaggregation by local segregation scores. The key property that these decompositions exploit is that $p_{u\cdot}^{t_2} = p_{u\cdot}^{t_1'}$, $p_{\cdot g}^{t_2} = p_{\cdot g}^{t_1'}$, $p_{u\cdot}^{t_1} = p_{u\cdot}^{t_2'}$, and $p_{\cdot g}^{t_1} = p_{\cdot g}^{t_2'}$, i.e. the equivalence of the margins. We can thus write:

$$\begin{aligned}\Delta_{\text{structural}} =& \frac{1}{2}(M(t_2) - M(t_1')) + \frac{1}{2}(M(t_2') - M(t_1)) \\ &= \sum_{u=1}^{U}\left(\frac{1}{2}p_{u\cdot}^{t_2}\left[L_u(t_2) - L_u(t_1')\right] + \frac{1}{2}p_{u\cdot}^{t_1}\left[L_u(t_2') - L_u(t_1)\right]\right) \\ &= \sum_{u=1}^{U}\Delta_{u,\text{structural}}\end{aligned} \tag{2.6}$$

Table 2.2: Decomposition of structural changes into contributions of each occupation

Occupation	Proportion		Observed		Counterfactual		Weighted difference
u	$p_{u\cdot}^{t_1}$	$p_{u\cdot}^{t_2}$	$L_u(t_1)$	$L_u(t_2)$	$L_u(t_2')$	$L_u(t_1')$	$\Delta_{u,\text{structural}}$
1	0.12	0.300	0.928	0.573	1.056	0.515	0.016
2	0.23	0.233	0.001	0.001	0.003	0.004	0.000
3	0.65	0.467	0.059	0.216	0.075	0.177	0.014

where $L_u(X)$ refers to the local segregation score for unit u in matrix X. The difference in structural segregation can thus be attributed solely to differences in the conditional probabilities, holding the marginals constant. Clearly, this decomposition is only possible because the M can be expressed as the weighted average of local scores. In the example, the decomposition results in three terms, one for each occupation. Table (2.2) shows the results for the detailed structural decomposition. Occupation one and three are responsible for the increase in structural segregation, while in occupation two, local segregation is low and almost unchanged. In more realistic settings with a greater number of units, the local segregation scores could now also be grouped by occupational major group or another characteristic (e.g., wage levels of occupations), if individual occupations are not of much interest. The sources of an increase or decrease in structural segregation, net of any marginal confounding, can thus be precisely understood.

2.4.4 Appearance and disappearance of units

Until now, it was assumed that at both points in time, all units and groups have non-zero counts. However, this assumption is often not met in practice. In the case of school segregation, schools may have closed down and new schools may have opened. In the case of occupational segregation, some occupations may have vanished and new occupations have become established. Capitalizing on the decomposition properties of the M, the approach used here can be extended to account for the effects of adding or removing units and groups.

Assume the simple case that in a labor market of five occupations, two occupations become obsolete:

$$t_1 : \begin{bmatrix} 5 & 15 \\ 15 & 5 \\ 10 & 10 \\ 5 & 15 \\ 15 & 5 \end{bmatrix} \quad \rightarrow \quad t_2 : \begin{bmatrix} 8 & 23 \\ 23 & 8 \\ 19 & 19 \\ & \\ & \end{bmatrix}$$

In this scenario, the workers from the vanished occupations were distributed across the remaining occupations, so that there are still 50 men and women each. Between t_1 and t_2, the M declines from .105 to .076. Is this purely an effect of the workers being redistributed? Or were the occupations that vanished more segregated than the occupations that remained?

To answer this question, define the set $S = \{1, 2, 3\}$ for the three remaining occupations, and $D = \{4, 5\}$ for the occupations that vanish. The sets S and D define "super-units" that are composed of individual units, and the share p_D is the proportion of workers in set D at t_1. The goal is to decompose $M(t_1)$ into the contribution of the occupations that vanish and those that continue to exist, which can be done using the general form of the between-within decomposition of M (Mora and Ruiz-Castillo 2011). Total segregation thus equals the between-super-unit M plus the weighted M within the two matrices defined by the two super-units, i.e.

$$M(t_1) = \underbrace{M\begin{pmatrix} 30 & 30 \\ 20 & 20 \end{pmatrix}}_{\text{between vanished/remaining}} + p_D \underbrace{M\begin{pmatrix} 5 & 15 \\ 15 & 5 \end{pmatrix}}_{\text{within vanished}} + (1 - p_D) \underbrace{M\begin{pmatrix} 5 & 15 \\ 15 & 5 \\ 10 & 10 \end{pmatrix}}_{\text{within remaining}}$$

Then solve for the last M term, which we call M^*:

$$M^*(t_1) = M\begin{pmatrix} 5 & 15 \\ 15 & 5 \\ 10 & 10 \end{pmatrix} = \frac{1}{1 - p_D}\left[M(t_1) - M\begin{pmatrix} 30 & 30 \\ 20 & 20 \end{pmatrix} - p_D M\begin{pmatrix} 5 & 15 \\ 15 & 5 \end{pmatrix}\right]$$

$$= .087 = \frac{1}{.6}\,[.105 - 0 - .4 \times .131]$$

This expression summarizes the mechanical effect of dropping occupations on the M index. To arrive at the "reduced M" on the left-hand side, subtract from M all the sources of segregation that are due to the vanished occupations only, which consists of a "between" and a "within" term. The between term summarizes how strongly the gender composition of the vanished occupations deviates from the remaining occupations, in total, while the within term summarizes how much segregation there is *within* the vanished occupations. The division by $1 - p_D$ has the effect of scaling the other occupations' proportions upward.[6]

$M^*(t_1)$ will be larger than $M(t_1)$ when the occupations that vanish were less segregated compared to the remaining occupations, and will be smaller in the opposite case. In this case, removing occupations 4 and 5 from t_1 reduces the M from $M(t_1) = .105$ to $M^*(t_1) = .087$. The "reduced M" can now be compared to the situation at t_2 using the regular IPF method. The approach outlined here thus amounts simply to a comparison of only those units that overlap across time points. However, an advantage of the M, which neither the H nor other indices have, is that there is an intuitive interpretation for the "missing" units.

Applying the decomposition to the example above gives the following:

$$\begin{aligned} M(t_2) - M(t_1) &= \Delta_{\text{removals}} + \Delta_{\text{marginal}} + \Delta_{\text{structural}} \\ 0.076 - 0.105 &= -0.017 + -0.006 + -0.006 = -0.029 \end{aligned}$$

In total, about 60% of the decline in segregation can be attributed to the effect of removing occupations 4 and 5. The remaining decline is equally due to changes in the marginals and to structural changes.

For simplicity, the example was only concerned with the removal of units, but additional units, such as newly arising occupations, can be handled in exactly the same way.

2.4.5 Summary of decomposition approach

The full, five-term decomposition of change between two segregation indices is thus:

[6]This can be clearly seen by assuming that we drop one unsegregated occupation only (occupation 3 from t_1). Then, the expression simplifies to $M^* = \frac{M(t_1)}{1-p_D}$. This shows that the mechanical consequence on M when an unsegregated occupation vanishes depends only on the size of the occupation, p_D.

$$
\begin{aligned}
M(t_2) - M(t_1) = {} & \Delta_{\text{appearing}} + \Delta_{\text{disappearing}} \\
& + \Delta_{\text{marginal-units}} + \Delta_{\text{marginal-groups}} \\
& + \sum_{u \in t_1 \cap t_2} \Delta_{u,\text{structural}}
\end{aligned} \tag{2.7}
$$

For most segregation problems, equation 2.7 is the minimum that is required to robustly understand changes in segregation, because the possible sources of change may point in opposite directions. Large changes in the marginals may hide worsening segregation at the structural level, or improvements in structural segregation might be overwhelmed by changes in the marginals.

Often, it is also of interest to compare *several* points in time or across space, and not just two. In this case, one point can be set as the reference point, with the decomposition then comparing all other points to the focal point in time. In a time series of occupational segregation, the first or the last point are obvious candidates for the reference point, while in a ranking of occupational segregation by cities the city with the median occupational segregation could be a good candidate.

Note also that this procedure can be used to decompose any M index. Because the cross-sectional decomposition of an M index again yield M indices, their change can also be studied over time. For instance, when studying occupational segregation, one might be interested in the change not only in the total M, but also for the partial M indices that define segregation within major occupational groups. (This will be done in the example below.) The total M admits to the following decomposition, assuming K major groups:

$$
M = M_{\text{between}} + \sum_{k=1}^{K} p_k M_k, \tag{2.8}
$$

where M_{between} refers to the gender segregation between the occupational major groups, p_k is the proportion of major group k such that $\sum_k p_k = 1$, and M_k is the segregation within major group k. When change is observed over time, the $k+1$ M indices defined in this decomposition can then be studied using the procedure outlined here.

2.5 Example 1: Occupational segregation

To consider the practical value of the above, I study occupational gender segregation in the U.S. between 1990 and 2016. IPUMS provides harmonized occupational codings based on the 1990 Census occupational codes for this period (Ruggles et al. 2018). The sample has been selected to comprise the employed, civilian population aged 16-66 with non-missing occupations. The occupational codes for 1990 were grouped into nine major groups (see Table 2.3).

When comparing occupations over time, two problems arise. First, the degree to which fine-grained occupations are recorded changes over time, and this is often a problem induced by the harmonization efforts. For instance, "Sociologists" are not coded separately in 2000-2016, but are available as a separate code in 1990. Second, occupations may vanish or new occupations may appear. "Stenographers," for instance, are no longer coded in later years, and this is probably because they no longer exist as a recognizable occupation. In many cases, it is hard to distinguish whether the problem is one of harmonization or one of disappearing occupations. For the purpose of this example, we will make the simplifying assumption that the harmonized occupations that are coded in each year represent recognizable, established occupations.

2.5.1 Descriptive statistics and total segregation

Table 2.3 contains descriptive statistics by year. Panel A shows the number of unique occupations that are available in each year, along with the number of categories that appear and disappear in each year. Panels B and C document a well-known pattern of occupational change, both in terms of female labor force participation and in terms of a changing occupational distribution. Panel D shows that there is considerable heterogeneity in terms of female labor force participation across occupational groups, and a heterogeneous pattern of change. In most occupational groups, female labor force participation increased, while in the Administrative and Operators/Laborers major groups, the share of women declined.

Table 2.3: Descriptive statistics

	1990	2000	2010	2016
Sample size (in 1000)	5917	6542	1443	1441
A. Number of occupations				
Number of occupations	369	336	330	319
Appearing occupations		0	0	0
Disappearing occupations		33	6	11
B. Labor force participation (%)				
Female	46	47	48	48
C. Distribution of occupational major groups (%)				
Managerial	12	12	13	14
Professional	13	16	17	18
Technical	4	4	4	4
Sales	12	11	11	11
Administrative	16	16	14	13
Service	13	14	17	17
Farming, Forestry	2	2	2	2
Production, Craft	11	11	9	9
Operators, Laborers	16	14	12	12
D. Female labor force by major groups (%)				
Managerial	43	44	45	46
Professional	54	57	59	59
Technical	46	48	49	48
Sales	49	50	51	51
Administrative	78	74	72	70
Service	57	59	60	59
Farming, Forestry	17	18	17	19
Production, Craft	8	10	10	11
Operators, Laborers	27	25	20	20

We calculate the M and the H for the total labor force, as well as separately for each major occupational group. This is based on the decomposition of the M into between and within-cluster terms, as in Equation (2.8). In this case, the between group term measures the segregation that is induced by the major occupational groups alone, while the within terms measure the segregation of detailed occupations within each major group. Because the number of observations are in the millions, bootstrapped standard errors are negligible (<0.0005) and therefore not shown.

The results are shown in Figure 2.1. Overall gender segregation, shown in the top panel, declined by 15% from 1990 to 2016 for the H and the M.[7] In 1990, the H was at 31%, and declined to 26% by 2016. The between term also declined, which means that major occupational groupings are becoming less informative about gender composition over time. However, major occupational groupings account for a large amount of overall gender segregation (45% of total segregation in 1990 and 42% in 2016).

While overall segregation declined, the within-terms reveal some heterogeneity. In most major groups, gender segregation declined. In others, notably Farming and Forestry as well as Production and Craft occupations, gender segregation increased strongly. This heterogeneity suggests that it is worthwhile to study major groups separately.

2.5.2 Decomposition of change

Many segregation analyses would stop at this point. Using the decomposition properties of the M, as well as the decomposition of change developed in this paper, we can go further and explore the patterns in more detail. To simplify the analysis of change, we focus on the changes between 1990 and 2016, without considering the intermediate years. Because no new occupations appear over time in this example, the total difference of any M term is thus de-

[7] Because the M is sensitive to the number of categories, one might suspect that the higher gender segregation in 1990 is an artifact of measurement. The normalization of the H index corrects for the changing number of categories, as shown above. Either way, in this case the variation in the number of occupations is too small to matter: if we restrict the calculation to the 317 occupations that are available at all five points in time, the M and H values are within 1% of the values presented in Figure 2.1.

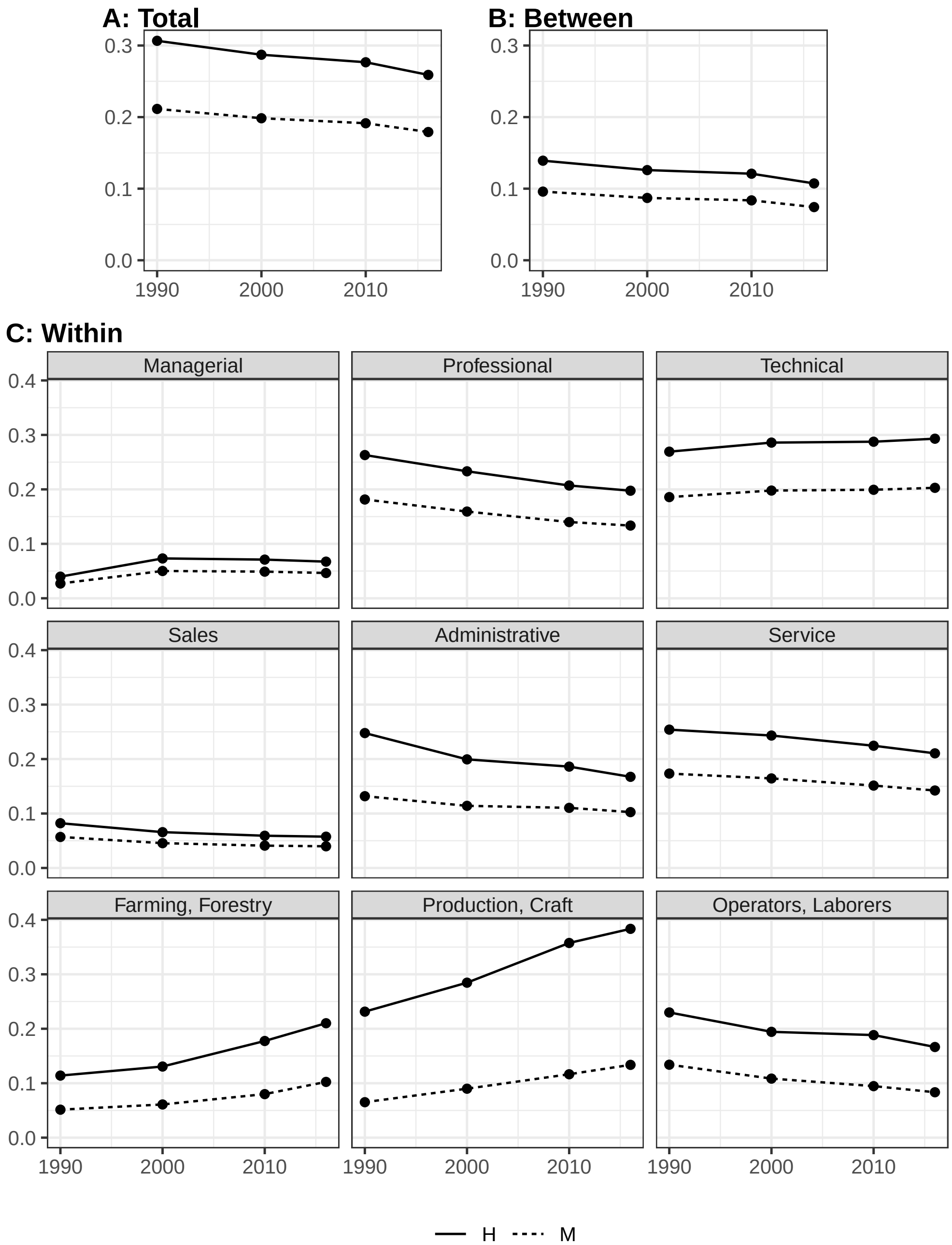

Figure 2.1: Occupational gender segregation, 1990-2016.

Note: Panel A shows total segregation by gender and detailed occupations. Panel B shows segregation between gender and major occupational groups. Panel C shows within-major-group gender segregation by detailed occupations.

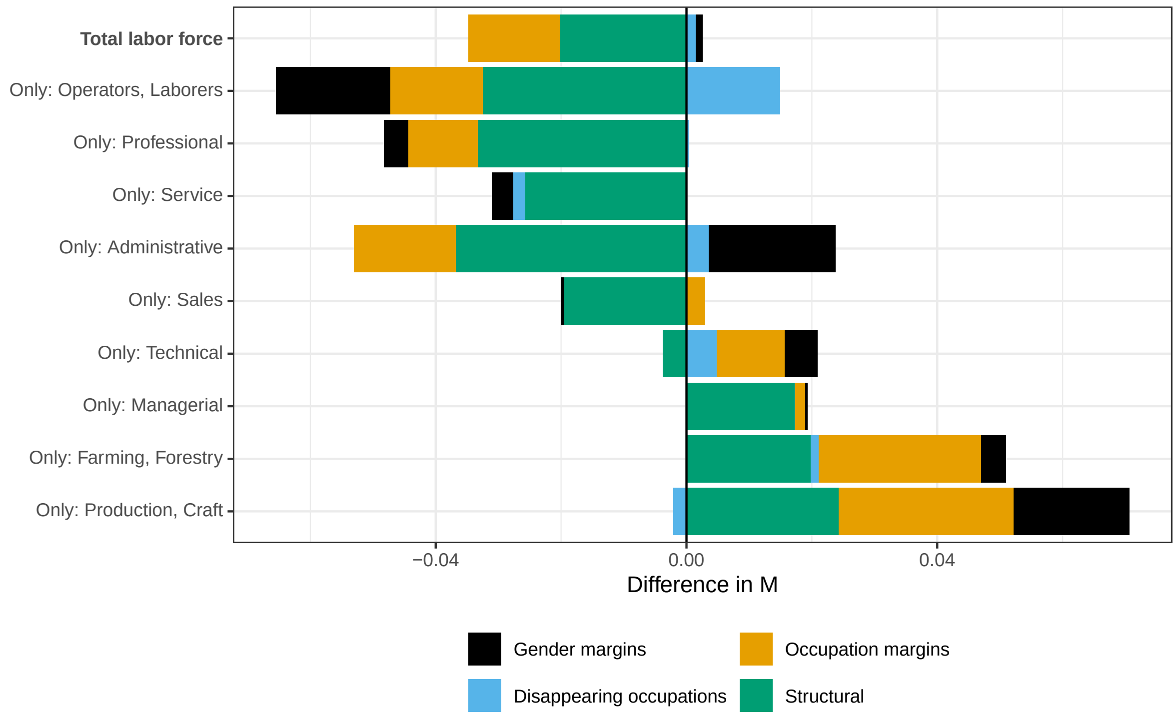

Figure 2.2: Decomposition of change

composed into four components: the effect of those occupations that are removed, the effect of the changing occupation marginal distribution, the effect of the changing gender marginal distribution, and the total structural component. Figure 2.2 shows the results graphically (without the between term), while Appendix B contains the full decomposition table. Again, the standard errors obtained through bootstrapping are negligible.

For the total M, the decline can be attributed to the changing occupational structure—i.e., the labor force has shifted towards occupations that are less segregated—, and, for the most part, to structural decrease. The decline in structural segregation accounts for 62% of the total decline in segregation. Most analysts of occupational segregation would consider this a positive development: Segregation decline is mostly due to declines in structural segregation, and the shift towards less segregated occupations has contributed even further to the decline. If all of the decline were due to the changing occupational margins only, we would still find that the average worker experiences less gender segregation. However, we could not conclude that the association of certain occupations with certain genders has lessened.

Segregation declined in five out of the nine major groups, and the share of the structural component was high in all five groups (between 65% and 117%). Within the major group of Operators and Laborers, the occupations that disappeared were relatively less segregated than the ones that remain, which increased segregation. However, the large marginal and structural components offset this small increase.

Segregation increased for four major groups. Except for the Managerial group, structural increase plays less of a role for these groups. For Technical occupations, structural change was in fact negative, but the marginal changes, especially the effect of the changing occupational distribution, led to an increase in segregation. In Farming and Forestry and Production and Craft occupations, structural segregation increased, but the changes in the marginals had a larger effect on the increase in segregation than the structural change. For the Managerial occupations, the increase in segregation is almost entirely due to a structural increase in segregation, which is worrisome. Overall, a rough pattern emerges: For those occupational major groups where segregation declined, it declined in large part because of a structural decrease in segregation. When segregation increased, it increased mostly because of changes in the marginal distributions—with the notable exception of Managerial occupations.

The increasing labor force participation of women accounts for only a minor part of the overall segregation difference: around 3% of the total change is explained by changing gender marginals. One might wonder why the sign of these effects does not correspond to the changing patterns of female labor force participation from Table 2.3. Shouldn't major groups in which women are rare show a decrease in segregation if the number of women increases? For instance, the female share of production and craft workers has increased from 8% to 11%, but this led to an expected *increase* in segregation. To understand why this is the case, consider the example of carpenters. In 1990, this occupation was 98.2% male, while the male share in the major group was 92.2%. This leads to a local segregation score for carpenters (within the major group) of $0.982 \cdot \log(\frac{0.982}{0.922}) + 0.018 \cdot \log(\frac{0.018}{0.078}) = 0.036$. In 2016, the share of male workers in the major group is 88.9%, which represents a reduction in the share of men of about 4% and an increase in the share of women of about 42%. After proportionally increasing the number of women and reducing the number of men, the *expected* share of carpenters that are men is now 97.4%. (To simplify, we only consider the forward adjustment here.) This leads to a counterfactual local segregation score for carpenters of $0.974 \cdot \log(\frac{0.974}{0.889}) + 0.026 \cdot \log(\frac{0.026}{0.111}) = 0.051$. This score is higher than before, although the number of women has increased. In this case, the expected effect of proportionally increasing the share of women within each occupation increases segregation, because it emphasizes existing patterns of segregation even more. The effect of the changing patterns of female labor force participation thus depends on the existing association structure between occupations and gender. This shows that the marginal effects have to be interpreted as *expected* changes in segregation when the odds ratios are held constant.

2.5.3 Comparison with other indices

IPF makes it possible to create a time series of adjusted M indices that is not confounded by marginal changes. To do this, we choose 1990 as the reference year and adjust the other years (2000, 2010, 2016) towards the marginals of the year 1990. Alongside with the adjusted M index, we also calculate the observed M and H indices, and the three other indices discussed above (see Appendix A for formulae).

The results for the five indices are shown in Figure 2.3. To ease comparison across the indices, the absolute numbers are transformed to be percentages of the 1990 values. First, it should be noted that all indices register a decline in segregation (although the A and SSD indices increase between 2000 and 2010). The structural decline, as calculated by the adjusted M, amounts to 10 percentage points of the 1990 value. The observed M and H "overstate" the decline, similarly to the V index. As seen in Figure 2.2, this is because the change in the occupational margins contributed to the decrease in segregation. Although the H is standardized, it gives essentially the same answer as the M. This is because the H is standardized by the gender distribution, which, however, had only a slight effect on segregation change. The effect of the occupation-margin-dependency of the H is thus clearly visible here. The other indices

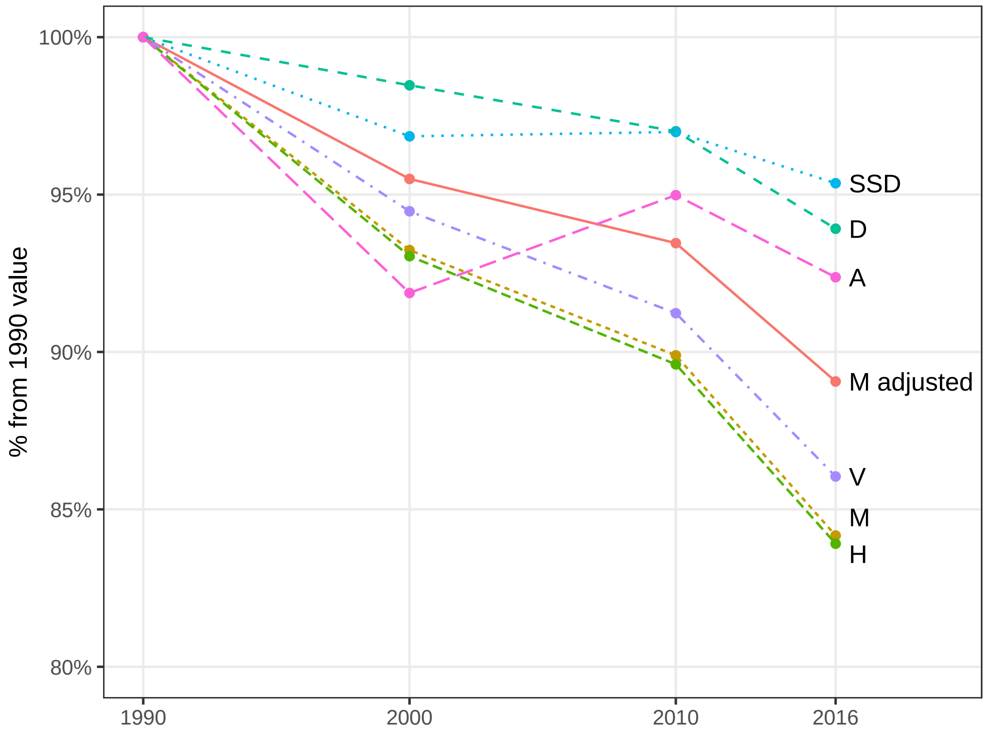

Figure 2.3: Comparison of the margins-adjusted M index with alternative indices

underestimate structural change compared to the adjusted M. The differences between the margin-free A and adjusted M are due to the different occupational weights. The A weights each occupation equally, which makes is susceptible to extreme values for small occupations that arise from sampling variability (Watts 1998), which is a possible explanation for its more erratic movement compared to the other indices.

The adjusted M index has a clear interpretation and a clear advantage: It quantifies the amount of segregation that is purely due to changes in the odds-ratios, net of any changes in the marginal distributions. It should be emphasized that the adjusted M is not a new segregation index, but just a regular M index, calculated on tables with identical margins. The main advantages of the decomposition will not be in the construction of an adjusted time-series, as in Figure 2.3, but in the ability to more precisely pinpoint where the changes in segregation originate.

Table 2.4: Decomposition of change

	Estimate
Index scores	
H in 2000	0.437
H in 2010	0.398
M in 2000	0.552
M in 2010	0.517
Difference in M	-0.035
Difference decomposition	
Additions	0.000
Removals	0.000
Racial group margins	0.000
Tract margins	0.003
Structural	-0.038

2.6 Example 2: Residential segregation

A second, short example illustrates the advantages of decomposing structural segregation. These results make use of the Longitudinal Tract Database (LTDB, Logan et al. 2014), which provides racial group counts for consistent Census tract boundaries. We just look at one example: The change in multigroup segregation in the borough of Brooklyn, New York City, from 2000 to 2010. Four racial groups are considered: Non-Hispanic Whites, non-Hispanic Blacks, Hispanics, and Asians.

Table 2.4 shows estimates of segregation by Census tracts in Brooklyn in 2000 and 2010, as well as the decomposition. The H declined from .437 to .398, which represents a decrease in segregation of about 9%. The difference in M values is then decomposed into the usual five terms. The main finding of this decomposition is that the decline in segregation is almost entirely due to structural change.

As a next step, the structural term is decomposed further to explore whether the declines in segregation are spatially clustered. We could use the terms $\Delta_{u,\text{structural}}$, as introduced in Equation 2.6, but these terms are weighted by tract proportion. To show changes at the scale of the

M index, we define instead the term ΔL_u which is just the average change in local segregation scores, net of marginal changes:

$$\Delta L_u = \frac{1}{2}\left[L_u(t_2) - L_u(t_1')\right] + \frac{1}{2}\left[L_u(t_2') - L_u(t_1)\right]$$

(This is simply Equation 2.6, with the weights $p_{u\cdot}^{t_1}$ and $p_{u\cdot}^{t_2}$ dropped.) Recall that the local segregation scores are measuring how strongly each tract's racial group distribution deviates from Brooklyn's overall racial group distribution. If a tract has exactly the same racial group distribution as Brookyln, its local segregation score will be 0; if a tract's racial group distribution deviates from Brooklyn's racial group distribution, local segregation for that tract will be > 0.

Figure 2.4 shows a map of Brooklyn, with the tracts shaded according to the value of ΔL_u, i.e. expected difference in local segregation when the margins are held constant. As Table 2.4 has shown, the average structural decline in segregation was ≈ -0.04. Thus, if all tracts were affected in the same way by structural segregation, we would expect ΔL_u to be -0.04 for all neighborhoods. The map shows that this is clearly not the case. Instead, declines in structural segregation have been much more pronounced in some neighborhoods of central Brooklyn, such as Clinton Hill, Williamsburg, or Bedford-Stuyvesant, which are shaded in dark blue. In some eastern parts of Brooklyn (Canarsie and East New York), as well as southwest of Prospect Park (the area of Sunset Park), structural segregation has increased, often quite strongly. Note that these values can be interpreted at the scale of the M. Thus, an increase in structural segregation of 0.2 for the whole of Brooklyn would mean an increase in segregation of about 36%, given the baseline value of $M = .552$ in 2010. This shows that the differences we observe across tracts are quite substantial.

This analysis could now be continued in various ways. As the index was calculated as a multigroup index, a further analysis might be interested in racial group differences. Another approach is to correlate the changes in structural segregation with tract-level measures, such as income or racial composition. It seems that segregation declined most strongly in gentri-

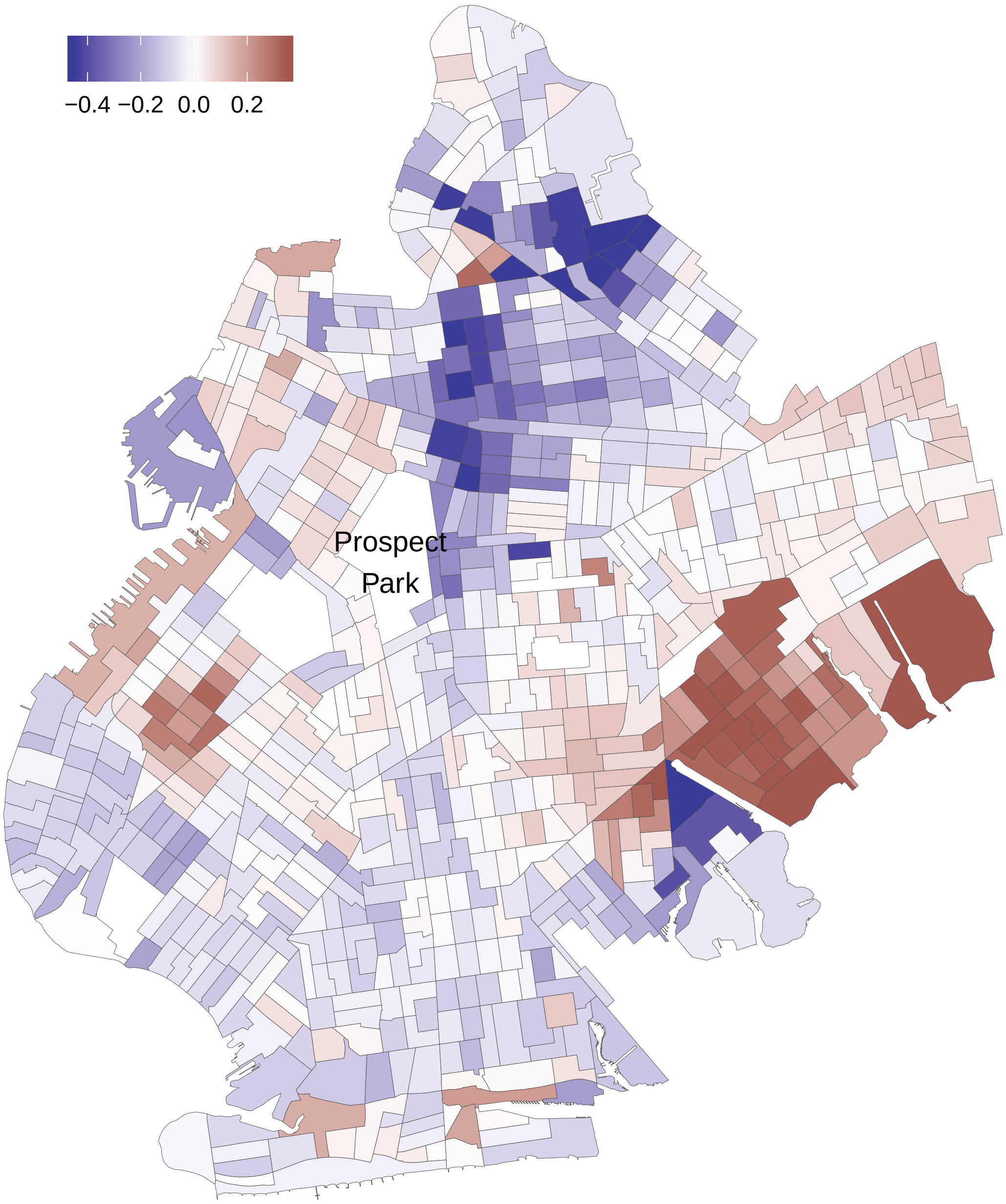

Figure 2.4: Tract-level differences in local segregation change ΔL_u, net of marginal changes

fying neighborhoods, while segregation has increased especially in the eastern, disadvantaged neighborhoods.

2.7 Limitations

The major limitation of the M index is that it is not standardized between zero and one. This clearly is a disadvantage. However, as has been pointed out throughout the paper, the full decomposition of change is only possible with the M index, as (a) it is decomposable into a weighted average of local segregation scores, (b) "vanished" and "new" units have a clear interpretation, and (c) the symmetry of the decomposition requires that the index is neither standardized in terms of groups nor in terms of units (if that were the case, the respective marginal component would be underestimated). In practice, one might therefore prefer to use the H index to establish the absolute level of segregation, and report all M changes in terms of percentages. This has been done throughout the examples.

A limitation of the decomposition method is its relative complexity (certainly compared to a computation of a time-series of segregation indices). This can be remedied through the use of the R package. Even with large tables and bootstrapping the computation of the decomposition will be fast.

In the segregation literature, there has been some concern about segregation indices that are calculated on the basis of small unit sizes or small group proportions. For instance, Winship (1977) derived expressions for the expectation of the index of dissimilarity for a city with two racial groups. With 10 households per block and varying proportions of the racial groups, the expected value of the D under a random housing pattern will range from .246 to .387. This represents serious bias. For the M and the H, the expected values[8] for the same situation range from approx. 0.053 to 0.058, and from 0.076 to 0.178, respectively, which is an improvement (see also Fossett 2017, p. 257ff.). The reason for this improvement can be seen when the M is expressed in terms of the individual table cells. In this formulation, the observed value in

[8]These values have been simulated using the "mutual_expected" function of the R package.

each cell, p_{ug}, is compared to the expected value under independence (by multiplying the two marginal probabilities, $p_{\cdot g}$ and $p_{u\cdot}$):[9]

$$M = \sum_u \sum_g p_{ug} \log\left(\frac{p_{ug}}{p_{\cdot g} p_{u\cdot}}\right) \tag{2.9}$$

Clearly, if p_{ug} is especially small, the logged ratio may be overly large. However, the expression is then weighted by p_{ug}, which leads to a relative decrease of the influence of the large ratio.

More generally, if one is concerned that in the problem at hand there may be zero segregation, and/or one deals with small group proportions or small unit sizes, one can take two steps to help remedy this problem: First, one can resort to the toolkit of classical statistics, such as Fisher's exact test or a chi-squared test. If these tests do not reject the null hypothesis of zero association between groups and units, then one can also conclude that there is no segregation. Second, one can use the observed marginal distributions to simulate random contingency tables, and compute the average segregation score for these tables. If the average simulated segregation score is > 0, the observed segregation score should be interpreted with caution. As a remedy, one could then combine units to arrive at a smaller contingency table. To check segregation bias for the H and M easily, the procedure has been implemented in the R package.

Lastly, while the literature has devoted considerable effort to "purge" the influence of marginal differences from segregation indices, it should be noted that differences in the marginal distributions may often be the relevant social fact compared to differences in structural segregation. Consider the comparison of a labor market over time. At the beginning of the period, men comprise 80% of the labor market , and at the end of the period the labor market has a balanced gender composition. One could compare the absolute level of segregation over time, but it seems that in such an extreme case, the relevant difference lies in the starkly different demographic profiles. When there are large changes in the marginal distributions (as in this case), it is also questionable whether the marginal and structural changes can be

[9]This equation also shows that the M is a "rescaled likelihood ratio test" (Card et al. 2013, p. 983), and provides a natural way to relate the M index to other approaches that are concerned with the study of association in contingency tables.

interpreted as (causally) independent. The IPF method would adjust the majority gender distribution at the first time point towards the balanced situation at the second time point, assuming that the marginal changes did not affect structural change (or vice versa). The IPF method would still sucessfully calculate the contributions of marginal and structural changes to this trend. However, if it were true that marginal changes causally produced all structural change, then the contributions from the (unknown) true causal model would be different. The deeper point here is that changing marginal distributions can be an important part of segregative processes, and that the summaries provided by a standardized segregation index should be used with caution when the margins are very different. An important empirical question to address in future work is how marginal changes and structural changes interact.

2.8 Conclusion

The paper presented a general method to decompose *changes* in segregation levels. It has been shown that the difference between two M indices can be decomposed into marginal and structural changes, as well as into terms that account for the appearance and disappearance of units. Parts of the method are, in principle, applicable to any segregation index. However, the advantages of the M index became apparent when considering changing sets of units under study and when a further decomposition of structural terms is desired. The decomposition of the structural term into the contribution of individual units is especially useful, as it may reveal important heterogeneity in segregation change among the set of units. The change in structural segregation allows a more precise testing of hypotheses about the causes and effects of changing levels of segregation at the unit level, and this change will be net of any influence of the marginal distributions. The benefits of this approach have been illustrated in the two examples.

The method described here can be applied to a variety of problems. Given that the M is a multigroup index, no measures have to be taken to account for segregation problems with more than two groups. Thus, the decomposition can be applied to school and residential racial segregation (as in example 2) where the analysis extends beyond just the majority-minority group dichotomy. Other examples where the method might be usefully applied are workplace racial and gender segregation. As firms shut down and new firms are founded, these studies typically have to account for a changing distribution of units. Furthermore, it is likely that there are firm differences in the propensity to segregate by race and/or gender (e.g., large and small firms).

The examples in this paper are focused on comparisons over time, but the method applies equally to comparisons across space. One useful application would be for comparisons of occupational segregation across countries or cities. One might suspect that observed differences in occupational segregation between cities are often due to marginal changes. For instance, a city with a large production sector will likely have higher gender segregation than a city with an employment profile skewed towards service occupations. One might suspect that the differences in segregation are just a consequence of the differences in the marginal occupational distribution. If one compares many cities or countries with each other, it will often not be feasible to compute all pairwise comparisons. Similar to dummy coding in regression analysis, one could instead choose one city as the "reference category" and compute comparisons against this reference city. Alternatively, one could also pool all of the city-specific datasets to capture overall segregation, and then compute differences of each city compared to the overall average.

The M index and the decomposition can be applied to a much wider array of problems than are usually considered in segregation analysis. The M index, as any segregation index, is a measure of statistical association between two categorical variables, and could thus be usefully applied to variables other than gender, occupation, racial groups, schools or firms. For instance, the study of social mobility relates the parental class distribution to the class distribution of the children. It could prove insightful to apply entropy-based indices to this problem as well, as it would allow researchers to make statements about which classes contribute the most towards increases and decreases in social mobility.

2.9 Appendix

A. Formulae for alternative segregation indices

For the alternative indices, the two-group versions are used. Let N be the number of occupations, T the total number of workers, M (W) the total number of male (female) workers, T_i the number of workers in the i-th occupation, and M_i (W_i) the number of male (female) workers in the i-th occupation. The notation follows James and Taeuber (1985) and Weeden (1998). D is the index of dissimilarity, SSD the standardized index of dissimilarity, V the variance ratio index, and A the closed form of the log-linear index.

$$D = \frac{1}{2}\sum_{i=1}^{N} \mid \frac{W_i}{W} - \frac{M_i}{M} \mid$$

$$SSD = \frac{1}{2}\sum_{i=1}^{N} \mid \frac{W_i/T_i}{\sum_{i=1}^{N} W_i/T_i} - \frac{M_i/T_i}{\sum_{i=1}^{N} W_i/T_i} \mid$$

$$V = \sum_{i=1}^{N} \frac{T_i(M_i/T_i - M/T)^2}{M(1-M/T)}$$

$$A = \exp\left(\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\log\frac{W_i}{M_i} - \left(\frac{1}{N}\sum_{i=1}^{N}\log\frac{W_i}{M_i}\right)\right]^2\right\}^{\frac{1}{2}}\right)$$

B. Full decomposition of Example 1

Table 2.5: Decomposition of change

Component	M 1990	M 2016	M Diff.	Disappearing occupations	Marginal Occupation	Marginal Gender	Structural
Total	0.211	0.179	−0.032	0.001	−0.015	0.001	−0.020
			(100%)	(−5%)	(45%)	(−3%)	(63%)
Between major groups	0.096	0.074	−0.022	0.000	−0.017	0.002	−0.006
			(100%)	(−0%)	(79%)	(−8%)	(28%)
Within major groups (sorted by Diff.)							
Operators, Laborers	0.134	0.083	−0.051	0.015	−0.015	−0.018	−0.033
			(100%)	(−29%)	(29%)	(36%)	(64%)
Professional	0.181	0.134	−0.048	0.000	−0.011	−0.004	−0.033
			(100%)	(−1%)	(23%)	(8%)	(70%)
Service	0.173	0.142	−0.031	−0.002	−0.000	−0.003	−0.026
			(100%)	(6%)	(0%)	(11%)	(83%)
Administrative	0.132	0.103	−0.029	0.004	−0.016	0.020	−0.037
			(100%)	(−12%)	(56%)	(−69%)	(126%)
Sales	0.057	0.040	−0.017	0.000	0.003	−0.001	−0.020
			(100%)	(−0%)	(−17%)	(3%)	(114%)
Technical	0.186	0.203	0.017	0.005	0.011	0.005	−0.004
			(100%)	(28%)	(63%)	(31%)	(−22%)
Managerial	0.027	0.046	0.019	0.000	0.002	0.000	0.017
			(100%)	(0%)	(8%)	(2%)	(90%)
Farming, Forestry	0.051	0.102	0.051	0.001	0.026	0.004	0.020
			(100%)	(3%)	(51%)	(8%)	(39%)
Production, Craft	0.065	0.134	0.069	−0.002	0.028	0.018	0.024
			(100%)	(−3%)	(41%)	(27%)	(35%)

Chapter 3: Training Regimes and Skill Formation in France and Germany: An Analysis of Change between 1970 and 2010

3.1 Introduction

How do educational systems prepare students for the labor market?[1] This question has been the focus of social scientists and policy-makers for a long time. In describing how countries establish a link between school and work, social scientists often rely on ideal-types of skill formation systems. An important distinction is made between systems that emphasize vocational training, where work and study are combined to prepare a student for work in a specific occupation or set of closely related occupations, and systems that emphasize general training, where education remains largely school-based and skill acquisition is more general. A large body of literature relies on this distinction, for example in trying to explain cross-national variation in youth unemployment, occupational mobility, status of the first job, or the length of job search (e.g., Bol and Werfhorst 2013; Breen 2005), among other topics. The question of whether skill formation systems are adequately equipped to deal with rapidly changing skill requirements in today's labor markets (e.g., through automation or globalization) has also sparked policy debates in many countries.

One of earliest and influential studies in the school-to-work literature is Maurice, Sellier, and Silvestre's book *The Social Foundations of Industrial Power* (1986, MSS hereafter), in which they studied how French and German men made the transition from school to work. MSS concluded that the French and German skill formation regimes have radically different approaches

[1]This chapter is coauthored with Thijs Bol and Thomas DiPrete. A slightly adapted version was published as: Elbers, Benjamin, Thijs Bol, and Thomas DiPrete. 2021. "Training Regimes and Skill Formation in France and Germany: An Analysis of Change between 1970 and 2010." *Social Forces* 99(3): 1113–1145. doi: 10.1093/sf/soaa037.

in organizing the match between schools and labor markets. France is described as an *organizational space*, where workers are hired based on their level of general education. Germany, in contrast, is described as a *qualificational space*, where jobs are closely matched to workers with specific qualifications that they obtained within the educational system. In Germany, the linkage between the educational system and the labor market is tight, resulting in a skill-formation system where specialized educational credentials predict labor market standing relatively well. In France, on the other hand, this linkage is weaker, and educational qualifications hold less information about the labor market positions of workers: workers with the same educational background end up working in vastly different occupations. MSS influenced several scholars who used this distinction as an important component of their theoretical work on the macro-structure of educational systems (Allmendinger 1989; Hall and Soskice 2001; Kerckhoff 1995; Shavit and Müller 1998).

The extent to which the organizational-qualificational dimension adequately captures the major differences between educational systems, however, has been called into question by DiPrete et al. (2017). They introduced a new method for measuring school-to-work linkages, building on the idea that strong school-to-work linkages imply that occupations are educationally homogeneous. Surprisingly, the authors found that France and Germany differ less than expected in terms of how closely educational pathways and occupations are linked. When comparing similar educational pathways across the two countries (say, university-educated workers in health), the French skill formation system often performs equally well in leading graduates to occupations that are a good fit to their skills (see also Bol et al. 2019) as the German qualificational space. Their analysis illuminated the fact that educational systems can differ both on a compositional dimension (the distribution of students across educational tracks), and in terms of their structure of association—the strength of linkage.

This similarity between the French and German skill formation system leads to a puzzle: If MSS were correct, then one or both countries must have changed. This would call into the question the quite static treatment of skill-formation systems in the literature. A second explanation for the puzzle is the treatment of gender in MSS. Their results are based only on the male labor force, while DiPrete et al. (2017) studied the whole labor force. We explicitly study potential gender differences, as generalizations about the two skill-formation systems may not be warranted when a large part of the labor force is disregarded.

This paper addresses this puzzle and maps how skill formation systems change by providing an empirical description of the possibly changing patterns of school-to-work linkages. We built upon earlier work by DiPrete et al. (2017), but introduce the dimension of *time* by focusing on the evolution of school-to-work linkages in France and Germany between 1970 and 2010. This requires the use of a method that is able to distinguish between compositional and structural changes, as in the last decades both France and Germany have seen educational expansion and an increase in the labor market participation of women—processes that are likely associated with changes in school-to-work linkages.

3.2 Country-level ideal-types

The study of skill formation systems has usually focused on country-level differences. Studies have classified educational systems along a number of dimensions (Allmendinger 1989), and this work was often influenced by MSS's qualificational-organizational distinction. A key aspect of this distinction is the focus on vocational specificity. In their comparative project, Shavit and Müller (1998) studied *rates* of vocational education (among other system characteristics) and their implications for the pathways that lead from school to work. Their key argument is that cross-national variation in how educational systems prepare students for the labor market is largely based on how many students are enrolled in vocational education:

> [W]e distinguish between two ideal-typical regimes of school-to-work transitions, which, following Maurice, Sellier, and Silvestre, we label *qualificational* and *organizational spaces*. The qualificational space is characterized by a high rate of *specific vocational education*. More precisely, a large proportion of the graduating cohorts leave the educational system with specific skills and occupational identities. This is in contrast with organizational spaces where education is predominantly *academic* or *general*, and where occupational skills are learnt on the job or in courses taken after leaving school." (Müller and Shavit 1998, p. 9)

The consensus in the literature is that school-to-work linkages are stronger in skill-formation systems that rely on occupation-specific vocational education and fit the "qualificational space" ideal-type (e.g., Breen 2005; Müller and Gangl 2003; Wolbers 2007). Students leave the educational system well-prepared and find work in an appropriate occupation relatively quickly, often because they enter the labor market with relevant work experience. This also has consequences for career mobility, which is higher in organizational spaces (Haller et al. 1985; König and Müller 1986).

While this literature has established the importance of skill-formation systems for later-life outcomes, we argue that the identification of countries with specific ideal-typical skill-formation systems rests on assumptions that have rarely been tested empirically.

First, the approach implies the possibility of distinguishing vocational educational programs that provide specific skills from academic programs that provide general skills other through a credentialing requirement to work in specific occupations, as is typically the case with those occupations that are called "professions". However, it is rarely discussed *how* the skill-content of educational programs can be measured. For instance, the UNESCO manual for the International Standard Classification of Education (ISCED) says virtually nothing about how skill content of vocational education is measured (OECD/Eurostat/UNESCO Institute for Statistics 2015, p. 65).

Second, even if it were possible to distinguish vocational and general programs, it would be an open empirical question whether specific skills are in fact associated with stronger school-to-work linkages. There is an implicit assumption that all vocational programs channel graduates to specific occupations, and so if there are more vocational graduates, the system will be more "occupation-specific." This focus on rates of vocational education asserts a substantive focus on the country as the unit of analysis: Rates are computed for the country as a whole; some countries are said to have vocational educational systems (Germany), and others do not have such a system (France or the U.S.). Given the considerable variation that is observed *between* different educational programs within the same country, we argue that a more fine-grained study of skill-formation systems is required.

3.3 A structural approach

We study the strength of school-to-work pathways as well as cross-national differences in their composition. We call this "linkage approach" a structural approach because (a) it does not rely on classifying programs as vocational or general, and (b) it measures the "success" of educational programs empirically by studying how closely a program is linked to specific occupations. The actual content of educational programs is difficult to ascertain, but a central characteristic of a "successful" vocational program is that it leads to one or a small set of specific occupations. This characteristic is *jointly* determined by educational content (e.g., length and intensity of the program, school-based vs. work-based programs) and by legal, cultural, and social characteristics of labor markets (e.g., occupational closure via licensing or collective bargaining agreements). We argue that a continuous, data-based *measure* of the strength of education-occupation linkage is more illuminating and empirically plausible than a binary, theoretically motivated *classification* (vocational/academic).

When comparing linkage across countries or within countries over time, a complication arises. Consider that someone who studied medicine in university is linked to the labor market as strongly as a graduate of a car mechanic vocational program. Assume further that the countries differ in their educational composition: While both countries have the same proportion of car mechanic graduates, the first country has a higher proportion of medicine graduates. The impact on total linkage from these two educational categories will then be higher in the country with the higher proportion of medicine graduates. The difference between the two countries will be purely an effect of the different marginal distributions. On the other hand, the marginal distributions could be identical, but the linkage scores for different occupations could be different (say, in one country medicine graduates link strongly because they are almost all doctors, while in another country they are sometimes doctors and sometimes administrators or managers in hospitals). Thus, when comparing countries or the same country over time, there are always two potential sources of difference: differences in the marginal distributions of education and occupation, and differences in the patterns of association.

Differences in marginal distributions are important, but the “success” of a country’s skill-formation system lies in the effectiveness with which it can link educational systems and labor markets. Even if many German students are enrolled in vocational programs, this fact tells us very little about the extent to which they obtain occupation-specific education. Similarly, even if in France a much smaller proportion of workers takes a vocational degree, the extent to which that vocational degree channels school-leavers to the same occupation might be high. We thus talk about the *strength of the association* between specific educational outcomes and the occupational structure, and we use the term “structural linkage” for this component of total linkage. The linkage approach provides a clear contrast with existing studies that focus attention on compositional differences only, i.e. rates of vocational education.

3.4 Patterns of change

An important question in the literature on skill-formation systems is how they change. Neo-institutionalism has argued that national educational systems were becoming increasingly standardized and focused on providing general education (Benavot 1983). The empirical evidence for this claim was the decline in the rate of secondary vocational education. The comparative project of Müller and Shavit (1998) came to a different conclusion. To be sure, the authors of the country-specific analyses saw change taking place to a greater or lesser extent in the countries they studied: For instance, France is described as a country that went through a dramatic expansion of secondary schools in the 1950s and 1960s (Goux and Maurin 1998), while Germany is described as a country with very stable educational institutions (Müller et al. 1998). However, these changes do not point universally in the direction of a decline of vocational education, refuting the neo-institutionalists’ view of convergence. On the contrary, Müller and Shavit (1998) found that skill-formation systems are relatively ’fixed’ in time.

In this paper, we revisit this question of *change*. We study both change in the educational and occupational marginal distributions (what we call compositional difference or change) as well as change in the structure of association between educational programs and occupations (what we call structural difference or change). We have several reasons for suspecting that not only the marginal distributions changed in important ways over recent decades, but also that the structure of association may have changed.

First, DiPrete et al. (2017) have raised questions about the current conception of French-German educational differences. France at present has a smaller proportion of workers who were educated in programs that have tight linkage to the occupational distribution than does Germany. At the same time, many of France's educational programs have linkage that is as strong or stronger than in Germany. The differences between Germany and France, in other words, appear to involve both structural and compositional differences. However the study by DiPrete et al. (2017) did not involve historical data. It may thus be the case that either country has changed in the recent decades, and that the differences were more pronounced in the 1970s.

Second, from the contributions in Müller and Shavit (1998) and many other works studying the transformation of Western economies, we know that both the educational and the occupational distribution experienced considerable change in recent decades. In most countries, higher education has seen an enormous expansion, and many countries also saw an increase in vocational education. At the same time, economies have shifted towards the service sector, often accompanied by deindustrialization and the privatization of formerly state-owned companies. Educational expansion has likely increased school-to-work linkages in the aggregate, but the increasing supply of highly-educated workers may have led to a decrease in the structural part of linkage via processes related to "overeducation." Similarly, the shift towards the service sector, where jobs are potentially more loosely defined in terms of their educational requirements, may have also contributed to structural declines in linkage.

Lastly, two important but rarely-discussed aspects of MSS's study are that it focused only on large industrial firms and only on the male labor force. This latter point is also true for follow-up studies (e.g., Haller et al. 1985; König and Müller 1986). The point about gender was argued by Marry et al. (1998):

> The majority of these studies have only been concerned with men [...], and one could ask whether the label “societal” is really appropriate for characteristics measured only for the masculine half of society. (p. 356, own translation)

Given the increase in female labor force participation rates since the 1970s, this critique has become even more important. In our analyses, we thus pay special attention to gender differences in the articulation of school-to-work linkages. The ideal-typical distinction that is still dominant today has been based only on the male labor force, and so we do not know to what extent “organizational” and “qualificational” spaces also exist for the full labor force.

3.5 Analytical strategy

3.5.1 A measure of school-to-work linkage

Our interest lies in the strength of the association between the educational system and the labor market as a function of time, operationalized here by the *educational level* and the *field of study* on the one hand (using the ISCED scheme), and *current main occupations* (ISCO-88) on the other hand. We use “field of study” here in a broad sense, referring to all education that is specific to an area of work, including vocational schooling, university degrees, but also the German *Ausbildung* in the dual system. The main analytical unit is the combination of educational level and field of study (“level-field”).

To measure the strength of the link between education and occupation, we employ a multi-group segregation measure, as introduced by DiPrete et al. (2017). The link between the educational system and the labor market is tighter when educational qualifications predict the occupations of workers well. When school-to-work linkages are low, educational qualifications provide no information about the occupations of workers. To capture this idea, we use the Mutual Information Index M (Mora and Ruiz-Castillo 2011; Theil and Finizza 1971).

We define G as the set of combinations of E ISCED levels and F fields of study. G are thus the "level-fields," which reflect the specialized educational qualifications that are available. To calculate M, consider a $J \times G$ contingency table, where each cell counts the number of workers that have one of level-field G and one of occupation J. To simplify the notation, we transform the contingency table from a table of counts to a table of proportions (i.e., estimates of probabilities). Each cell's p_{jg} represents the proportion of workers in occupation j who were educated in level-field g, and therefore $\sum_j \sum_g p_{jg} = 1$. We also define the marginal probabilities $p_{\cdot g} = \sum_j p_{jg}$ and $p_{j\cdot} = \sum_g p_{jg}$, which simply reflect the overall level-field and occupational distributions. We define the conditional probability of being in occupation j, given a level-field g, as $p_{j|g}$.

We define $M(G; J)$ as "total linkage," which measures the dependency between the level-fields and occupations contained within G and J. Linkage is high for a specific level-field when the occupational distribution of that level-field deviates strongly from the overall occupational distribution. We call this the local linkage L of level-field combination g, and define it as

$$L_g = \sum_j p_{j|g} \log \frac{p_{j|g}}{p_{j\cdot}}. \tag{3.1}$$

Local linkage will be minimized at zero when $p_{j|g} = p_{j\cdot}$ for all j. In other words, a level-field is not linked to the labor market if the occupational distribution for workers that have this level-field is identical to the overall occupational distribution. The more strongly the occupational distribution for a level-field deviates from the overall occupational distribution, the higher the local linkage for this level-field. The possible maximum value for local linkage depends on the overall size of the smallest occupation. For instance, assume that the smallest occupation comprises 1% of the labor force. If all workers with level-field g worked in only this occupation (i.e. $p_{j|g}$ is 1 for one j, but 0 for all other j), then local linkage would be maximized at $\log \frac{1}{.01} \approx 4.6$.

To characterize overall linkage, M is defined simply as the weighted average of the local linkage scores for all level-fields G:

$$M(G; J) = \sum_g p_{\cdot g} L_g. \tag{3.2}$$

This additive decomposition of M is helpful in determining where in the educational distribution the linkage strength of a country originates. Further details can be found in the more extensive treatment by DiPrete et al. (2017).

3.5.2 Studying change

The M is not a margin-free measure of segregation, i.e., segregation may increase or decrease depending on changes in the marginal distribution of level-fields or occupations. This is immediately apparent from Eqs. 3.1 and 3.2. This means that changes in M over time can arise through different mechanisms: either because of a change in the margins without any structural changes, or through a change in the structural association between level-fields and occupations. Given that we are interested in studying differences over time and across countries, we seek to isolate marginal and structural changes.

We adopt a procedure first proposed by Karmel and Maclachlan (1988), which was extended and described in the previous chapter. This procedure is based on the insight that the odds ratios of the contingency table are the only measure of association that does not change when the margins are transformed. Consider two $G \times J$ contingency matrices for the same country at different points in time, t_1 and t_2. To make a margin-free comparison, we adjust the margins of the contingency matrix at t_1 to be identical to those at t_2. This is achieved by the use of iterative proportional fitting (IPF), where first the row margins of t_1 are scaled towards those of t_2, and then the column marginals of the resulting table are scaled towards those of t_2. Repeating this process several times will transform the marginals of t_1 to approach the marginals of t_2, while preserving the odds ratios of the matrix as of time t_1. The process is repeated until the marginals are within 0.001% of the marginals at t_2. We call the resulting, counterfactual matrix t_1'. We then repeat the IPF procedure, starting from matrix t_2, to arrive at matrix t_2'. These two scenarios differ in the choice of the original matrix: One adjusts from t_1 forward towards t_2, while the other adjusts in a backward direction from t_2 to t_1.

Given the four matrices, we calculate $M(t_1)$ and $M(t_2)$ as the observed linkage at time t_1 and t_2. We also compute $M(t_1')$, as the adjusted t_1 linkage, which can be regarded as the counterfactual linkage at time t_2 if only the margins had changed to equal their values at time t_2. Because $M(t_2)$ and $M(t_1')$ have the same marginal distributions, they differ only in their association structure (i.e., the odds ratios). The same goes for $M(t_1)$ and $M(t_2')$. To arrive at a single estimate, we average the forward and backward scenarios (often called a Shapley decomposition):

$$\begin{aligned} M(t_2) - M(t_1) = & \overbrace{\frac{1}{2}(M(t_2) - M(t_2')) + \frac{1}{2}(M(t_1') - M(t_1)}^{\Delta_{\text{marginal}}} \\ & + \underbrace{\frac{1}{2}(M(t_2) - M(t_1')) + \frac{1}{2}(M(t_2') - M(t_1))}_{\Delta_{\text{structural}}} \end{aligned} \tag{3.3}$$

Another benefit of this approach is that it allows a further decomposition of the structural and the marginal component. The marginal change can be subdivided into two components: one component quantifies the contribution of changing educational marginals and one quantifies the contribution of changing occupational marginals. The structural component can be decomposed into the contributions of each individual level-field. We make use of this property by summing the contributions of the level-fields using ISCED categories and fields of study.

The method described here can be used to compare any two M measures, where "t_1" and "t_2" can stand either for different points in time or for different countries at the same time.

3.5.3 Data

The European Labor Force Survey (EU-LFS, Eurostat n.d.) is well suited for our purposes because it provides harmonized variables for educational levels, fields of study, and occupations. In the 2005 to 2010 datasets, educational levels and fields of study are coded in the ISCED-1997 scheme, and occupations are coded using ISCO-88 (3 digits). In 2011, the EU-LFS switched to ISCO-08, which is why we use only the years up until 2010. Because of the small sample size in Germany, we pool the samples for 2006-2007 and for 2008-2009. In the EU-LFS, no fields of study were recorded before 2005. For France, we additionally use the “Formation et Qualification Professionnelle” (FQP, INSEE/ADISP-CMH n.d.(b)) survey for 1970 and 1985, as well as the 1990-2002 series of the French Labor Force Survey (Enquête Emploi, INSEE/ADISP-CMH n.d.(a)). Here, fields of study are recorded consistently beginning in 1995. To increase the sample size, we pooled the years 1995-1997, 1998-2000, and 2001-2002. For West-Germany, we use the Public Use Files of the 1970 and 1987 censuses (RDC of the Federal Statistical Office and Statistical Offices of the Länder n.d.).[2] We restricted all samples to the current active workforce aged 15 to 64[3], leaving out students and the unemployed.

Educational level, field of study, and current main occupation were recorded in different schemes in many of these years. For educational levels, we used the official ISCED mappings to code degrees into the ISCED-1997 scheme (see Table 3.1). The EU-LFS does not provide a breakdown for category *3ab*, which combines workers with general and specialized education. We therefore split this category using the field of study information: We code a worker as *3ab_voc* if a worker completed a professional or technological *baccalauréat* (maturity exam) in France or an *Ausbildung* (dual training) in Germany; we code as *3ab_gen* for the general *baccalauréat* in France and the *Abitur* (maturity exam) in Germany. We merge *5a* and *6* (PhD) because the latter is a small category and cannot be distinguished in earlier surveys. We also merge ISCED 0 and 1, because these two categories can not be distinguished in earlier years. In categories *1*, *2* and *3ab_gen*, workers have obtained general education, while in all other categories workers have obtained specialized education (indicated by the presence of a field of study), be it either through dual training, vocational schooling, or in higher education. Programs classified as *5b* are advanced vocational programs on the verge of tertiary education, while programs in *3c* and *3ab_voc* are more often a mix of on-the-job training and school-based training.

Fields of study have been manually coded into ISCED fields of study (see Table 3.2). The crosswalks are found in Appendix B. Except for the FQP survey in 1970, we were able to use

[2]The Mikrozensus cannot be used for our purposes, because up until 2004 it is missing the field of study for people who were trained in the dual system.

[3]We examined other age cutoffs (Appendix D).

Table 3.1: ISCED-1997 and native degrees

Category	**France**	**Germany**
ISCED 1 or less	Elementary education or less	Elementary education or less
ISCED 2	Brevet, BEPC or some secondary education	Hauptschul-/Realschulabschluss
ISCED 3c	CAP, BEP, BP	
ISCED 3ab_voc	Baccalauréat professionnel Baccalauréat technologique	Lehrabschluss, short vocational school
ISCED 3ab_gen	Baccalauréat général	Abitur, Fachhochschulreife
ISCED 4		Abitur/Fachhochschulreife *and* Lehrabschluss
ISCED 5b	DUT, BTS infirmier, assistante sociale	Meister, Techniker, long vocational school (Fachschule, e.g. in health)
ISCED 5a/6	DEUG, License, Maîtrise, Diplôme, etc. DEA, Diplôme de docteur	BA, MA, Diplom, Magister, etc. Promotion

Note: Adapted from official ISCED mappings for France and Germany (`http://uis.unesco.org/en/isced-mappings`) proportional crosswalks to harmonize native occupational codings into ISCO-88 (one of these crosswalks is constructed using data provided by BIBB 2006). To create a proportional crosswalk, we identified surveys where the native occupational scheme and ISCO-88 were coded for the same individuals. For each native code, we then calculated the proportion of double-coded ISCO-88 codes. For instance, the French PCS-1982 code 3751 (*Cadres de l'hôtellerie et de la restauration*) is coded in ISCO-88 as 122 (Production and operations department managers) in 58% of the cases, and 131 (General managers) in 42% of the cases. When we apply the crosswalk to our data, we randomly choose with a probability of .58 the ISCO code 122, and 131 otherwise. This process introduces uncertainties into our estimates, which are in practice very small (see Appendix C).

3.6 Results

3.6.1 Descriptives

Table 3.2 provides descriptive statistics by year. Percentages are reported for gender and age groups, as well as for our three main variables of interest.

The tables confirm the well-known patterns of the changing demographics of labor force participation. Women's labor force participation increased from 37% to 48% in France and from 36% to 46% in Germany—always slightly below the French levels. The changes in the age structure reflect both an aging population and an increase in educational attainment. Many of those aged 15-24 in 1970 were already in the labor force, while in 2010 many in this age group are still in vocational or tertiary education, entering the labor market at older ages.

Of special interest are the dramatic changes in the educational structure. In Table 2.3, the effects of educational expansion are clearly visible, with increases in vocational (ISCED *3ab_voc*, *3c*, and *4*) and tertiary education and strong declines in the share of workers with a "general" field, i.e. ISCED levels *1*, *2*, and *3ab_gen*. For both countries most of the expansion of tertiary education was in higher tertiary education (university and post-graduate degrees, ISCED *5a/6*) instead of lower tertiary courses (ISCED *5b*).

Table 3.2: Descriptive statistics by year

	Germany				France						
	Census		EU-LFS		FQP		Enq. Emploi			EU-LFS	
	1970	1987	2005	2010	1970	1985	1996	1999	2001	2005	2010
Sample size (in 1000)	1146	1261	185	20	27	28	67	67	45	54	77
Gender											
Female	36	38	45	46	39	42	44	45	45	46	47
Age											
15-24	17	15	6	6	21	13	5	5	6	7	7
25-34	26	26	20	20	21	30	28	28	26	26	25
35-44	23	24	33	28	24	27	31	30	30	29	29
45-54	18	25	27	30	20	20	27	29	29	27	27
55-64	15	10	13	16	14	10	9	8	10	10	12
Educational levels (ISCED)											
ISCED 1	0	1	2	2	39	20	13	10	9	9	6
ISCED 2	32	22	11	8	34	32	22	21	20	18	16
ISCED 3c					16	28	31	31	30	30	27
ISCED 3ab_voc	52	52	49	50	1	3	6	7	7	14	16
ISCED 3ab_gen	1	2	4	2	3	6	6	6	7	1	1
ISCED 4	1	2	7	8							
ISCED 5b	7	8	11	10	2	6	11	12	14	12	14
ISCED 5a/6	7	13	17	19	3	6	11	12	13	17	20
Share of general education only											
General/No field	33	25	17	13	77	58	41	38	36	28	24
Educational fields (ISCED)											
Teacher training, education	2	4	4	4	0	0	0	0	0	1	1
Humanities, languages, arts	2	2	4	4	6	5	4	4	5	8	8
Social sciences, business, law	29	30	30	30	24	28	32	33	33	32	34
Science, maths, computing	2	3	3	3	10	8	3	4	4	8	8
Engineering, manufacturing	50	46	36	34	42	38	38	37	35	31	29
Agriculture, veterinary	4	3	3	3	5	5	5	4	5	4	4
Health, welfare	6	8	13	14	6	11	11	11	11	10	10
Services	5	5	8	8	6	5	7	7	7	5	5
Occupation (ISCO-88, 1 digit)											
Armed forces	1	1	1	1	1	1	1	1	1	1	1
Managers/Senior Officials	3	5	6	6	8	7	8	8	8	8	9
Professionals	7	11	15	16	6	8	10	10	11	13	14
Technicians/Assoc. Profess.	17	21	22	22	10	15	17	17	18	18	19
Clerks	11	13	12	12	12	15	15	15	14	12	11
Service/shop workers	9	10	12	12	6	9	12	12	12	12	13
Skilled agricultural workers	5	3	2	2	14	8	5	4	4	4	3
Craft and related workers	25	20	15	14	20	16	14	14	13	12	11
Plant operators/assemblers	10	8	8	7	10	13	11	11	11	9	9
Elementary occupations	11	10	8	8	13	8	8	8	8	10	10

If we maintained a focus on *rates* of vocational education, then clearly the French skill-formation system has changed enormously: for instance, the combined share of graduates in *3c* and *3ab_voc* has risen from 17% to 43%. The numbers also show clear differences in the two countries' skill-formation systems. In 1970 Germany, roughly 60% of the labor force had undergone vocational education (ISCED *3ab_voc*, *4*, and *5b*), compared to less than 20% in France. In 2010, France's educational distribution resembles Germany's distribution in 1970. The relative shares of both vocational and tertiary education in France have increased enormously, leading to a labor force in 2010 with over three quarters of the workers having some sort of specialized education, up from less than one quarter in 1970. This underscores the stability of the German skill formation system compared to a number of educational reforms in France that aimed to increase vocational and tertiary education levels (Brauns et al. 1999; Day 2001). This pattern of stability in Germany and change in France is a recurring theme in this paper.

Vocational education is traditionally focused on manufacturing and business degrees, which in both countries constitute the majority of degrees awarded. With deindustrialization and an increasing focus on service occupations, the *relative* share of manufacturing degrees has decreased markedly over time in both countries. The increasing diversity of vocational education is reflected by an increase in health, welfare, and services degrees. When comparing the distribution of fields of study between the countries, they show a roughly similar pattern.

The distribution of occupational major groups and the patterns of occupational change are similar in both countries, suggesting changes in the skill distribution of the labor force that are common to both economies. The growth has been concentrated in high- and medium skilled occupations (groups 1 to 5), while lower-skilled occupations have declined (groups 6-9). There are also some important differences between France and Germany that seem stable over time. In both 1970 and 2010, Germany had more workers in craft and professional occupations (groups 2, 3, 7), while France had more workers in agriculture, low-skilled, and management occupations (groups 1, 6, 9). This pattern is consistent with MSS's findings about the organiza-

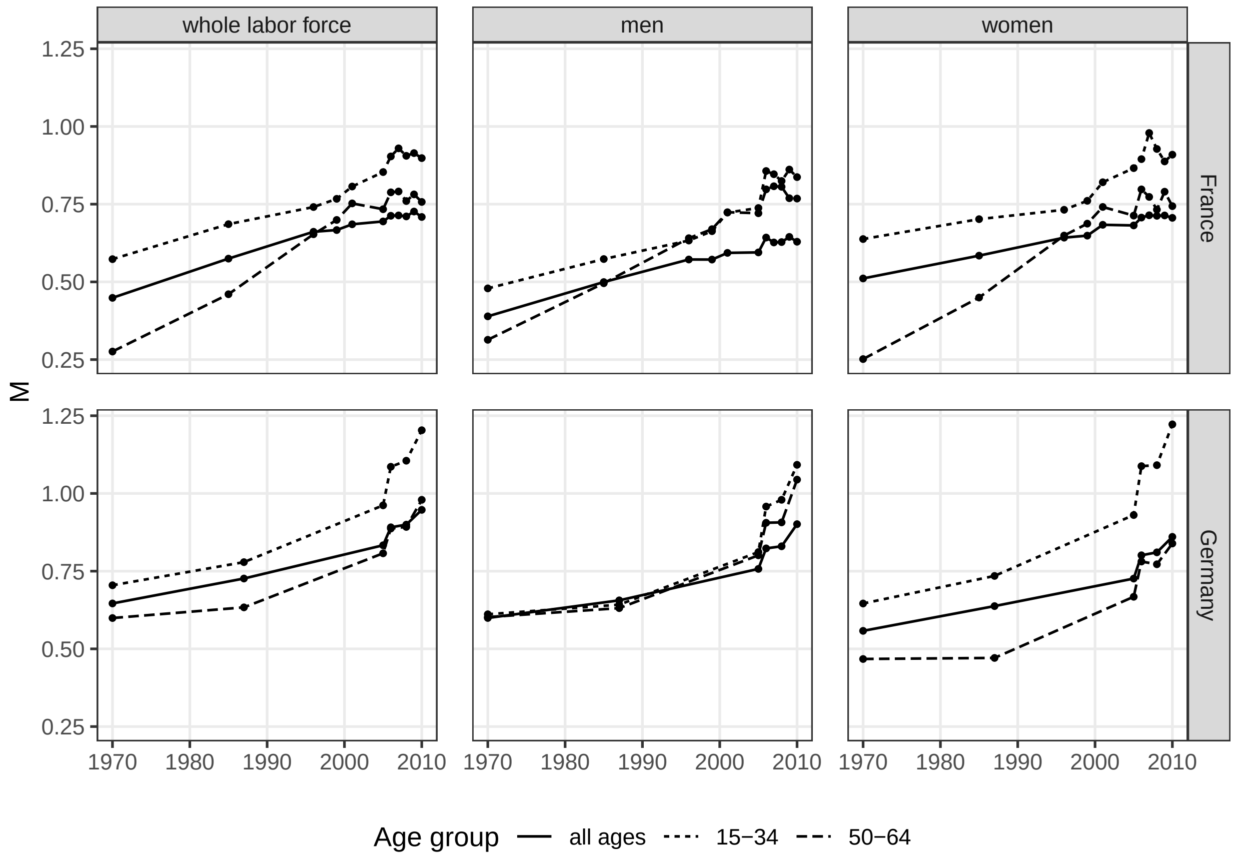

Figure 3.1: Aggregate linkage over time, by gender, age, and country

tion of work in the two countries. In France, the number of low-skilled workers and managers is higher, while Germany relies more on specialized, “medium-skilled” workers.

3.6.2 Observed total linkage over time

Figure 3.1 shows the strength of linkage for France and Germany both for the whole labor force, as well as for men and women separately. Each panel also contains a breakdown by younger (ages 15-34) and older workers (ages 50-64). The M is measured for each year and gender-age-country combination separately and defined by applying eq. 3.2 to the matrix of level-fields and three-digit ISCO-88 occupations.

The absolute value of the M is not strictly interpretable, as the values it can take depend on the dimensions of the contingency table. While an absolute interpretation is not possible, by harmonizing the data, we can compare the M scores in terms of their relative differences—countries over time, between countries over time, or between genders. For this reason we will interpret the changes in M as percentage differences.

Concentrating first on the whole labor force, Germany's 2010 M was about 38% higher than France's. This is similar to the difference reported in DiPrete et al. (2017), who used different datasets. In 1970, Germany's M was 50% higher than France's. Since 1970, Germany's overall M has increased by 56%,[4] and France's M by 70%, with most of this change occurring before 2000. In terms of total linkage, the countries have thus slightly converged over time.

The gender breakdown reveals that in both countries, the rate of linkage change has been similar for men and women. However, in Germany, men have slightly higher linkage than women, while in France the opposite is true. One reason for this could be the different gender-specific occupational composition of the French and German labor markets. A large share of the German male workforce is employed in the skilled trades and crafts, which tend to be well-linked. In France, in contrast, men are more often employed in clerical and service occupations that tend to be linked less strongly.[5]

A more striking pattern emerges when one takes into account the differences between age groups. Figure 3.1 shows that age differences in linkage are small for German men, while they are much larger for German women. Age group differences are larger for men in France than in Germany, but in France (just as in Germany), the age differences in linkage strength are stronger for women than for men. To put it differently, older women are in occupations that link less well than is the case for younger women in both countries, and this age gap is bigger for women than for men. However, the two countries differ in the pattern of change. In France, a visible convergence of linkage strength for older and younger workers has occurred in recent years

[4] Germany's M showed an especially strong increase between 2005 and 2010. We discuss possible reasons for this, as well as possible consequences for our later results in Appendix E.

[5] We emphasize that the analysis only takes into account the employed population. Future research might address the question of how changing labor force participation rates by education have influenced school-to-work linkages, especially for women. We present a short analysis in Appendix F.

for both men and women. In contrast, the linkage-strength gap between older and younger German women has not closed over the forty years covered by our data.

Finally, while the size of the age gap varies by country and gender, the relative pattern is similar and in the expected direction: Younger workers who have just begun their careers link more strongly than older workers with notably stronger age effects for women than for men. The small differences in linkage strength for older and younger German men suggests a relative stability of educational institutions across cohorts and a relatively minor impact of career mobility for the linkage strength of German male workers. In France, the temporal gains in linkage strength for younger workers may signal that educational reforms have shifted educational attainment towards more strongly-linking educational degrees. The even stronger upward shift for older male French workers may reflect the lagged effects of shifts for younger workers combined with career mobility patterns that produce higher linkage strength. The especially strong temporal gain in linkage strength for older female French workers may reflect changing implications of motherhood on their composition and occupational placement as well as lagged early career effects.

A pattern that is not easy to distill from Figure 3.1 is that cross-national variation in linkage highly depends on which gender one focuses on. Figure 3.2 sheds light on this issue and shows the difference in linkage between France and Germany by gender. When only taking men into account (Figure 3.2, left panel), our results seem to support MSS's main conclusions: Germany provides a tighter school-to-work linkage that France. This conclusion is starkly different for women (right panel). Especially during the 1970s (the likely period of MSS's research), linkage for women is only slightly higher in Germany than it is in France and the similarities between the two countries are more striking than the differences. Combined with the results from Figure 3.1, our findings question a clear separation between the German and French skill formation systems. Are they indeed the two extremes of the ideal-typical distinction between qualificational and organizational spaces, or is this conclusion by MSS only applicable to part of the labor force?

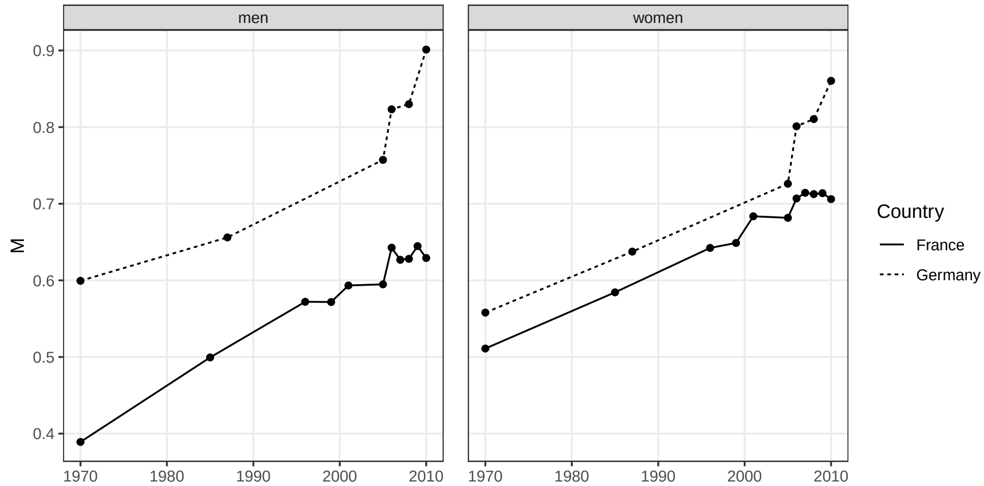

Figure 3.2: Gender-specific patterns of linkage

Next, we use eq. 3.2 to study where the linkage strength in Germany and France originates. The full decomposition of local linkage scores gives 35 terms, one for each level-field. For a more parsimonious presentation, we sum the contribution of level-fields to the total M by simplified ISCED levels, where we group tertiary education (ISCED *5a/6*), upper vocational/lower tertiary education (*5b*), vocational education (*3ab_voc, 3c, 4*), and general education (*1, 2, 3ab_gen*). The four components are plotted separately by country-years in Figure 3.3. The percentage indicates the relative contribution towards total linkage strength, and it is instructive to compare this number to the proportion of the respective level among the labor force from Table 3.2.

In Germany, almost half of the total linkage strength in 1970 originates from upper secondary vocational education (ISCED levels *3ab_voc/3c/4*). Although only 7% of graduates had obtained tertiary education (*5a/6*), this component accounts for 27% of total linkage strength. General education (i.e., ISCED *1/2/3ab_gen*) comprised one third of Germany's labor force in 1970, but contributes only 13% to total linkage. In 2010, these relative contributions have changed only little, and seem to be mostly explained by changes in the relative proportions. France in 1970 was dominated by general education: 78% of the labor force had a general degree, and this group contributes 34% towards total linkage. While the other three components together constitute only a small share of the labor force, they contribute a considerable amount.

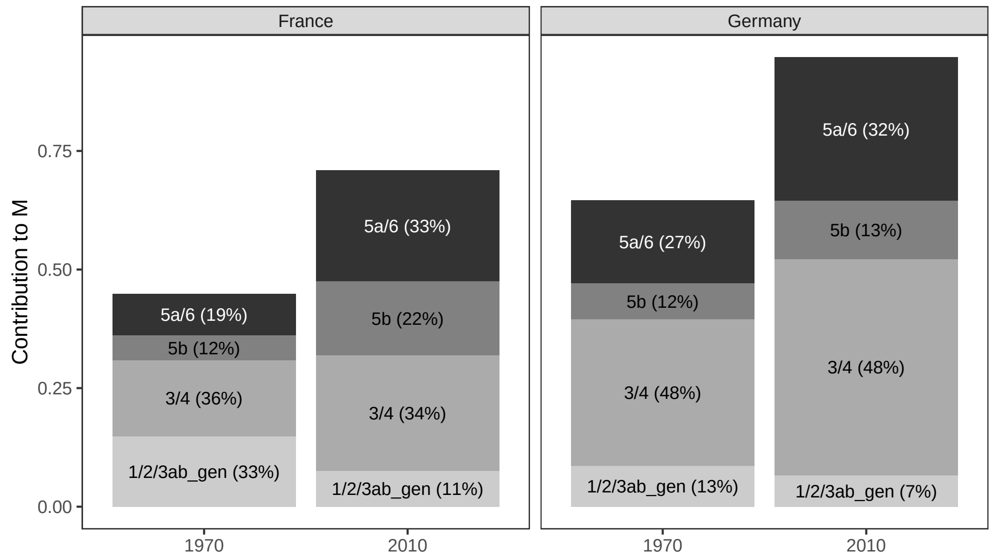

Figure 3.3: Decomposition of M by ISCED levels
Note: ISCED 3/4 refers to ISCED 3ab_voc, 3c, and 4.

Two thirds of France's linkage in 1970 originates from just 22% of the labor force. France's changing educational distribution is also reflected in a different pattern of linkage: In 2010, general education contributes much less to linkage, and the contributions of ISCED 5 and 6 have grown considerably.

3.6.3 Between-country differences: then and now

As we noted above, stratification research has often equated high *rates* of vocational training with high vocational specificity and treated this as a structural feature of the country's educational system. Figure 3.3 already shows that this conclusion might not be justified: In France in 1970, vocational graduates were only a small share of the labor force, but this group linked strongly to the labor market. Thus, the internal structure of school-to-work linkages in France and Germany does not match the conventional wisdom about these two countries.

The findings presented so far have to be interpreted with caution, as they do not fully account for differences in the educational and occupational marginal distributions. Comparisons between countries and over time can be misleading when approached through Figure 3.3, because they combine structural and marginal differences. We now present decompositions that allow us to disentangle the possible sources of difference.

First, we apply eq. (3.3) to the differences *between* countries, both in 1970 and in 2010.[6] (t_1 stands for Germany, and t_2 for France). The results are shown graphically in Figure 3.4, where the left panel decomposes the differences for the complete labor force, and the right panel contains results separately by gender. A negative score means that the French component is lower than the German component.

The top panel shows the raw difference in linkage between the two countries as observed in Figure 3.1. This total difference is decomposed into the three components shown in the second panel from the top (titled "Overall"): "occupational margins," "educational margins," and "structural." As we have described above, the structural change can be further decomposed into the contributions of the 35 level-fields. For an interpretable presentation of the results, we chose to sum those 35 contributions once by level, and once by field. These two decompositions are presented in the third and fourth panel from the top, titled "Structural: Levels" and "Structural: Fields."

Focusing first on the overall decomposition, France's linkage was lower than Germany's in 1970, but this difference is almost completely accounted for by differences in the educational marginal contributions. This means that the stark differences between Germany and France that we observe are mostly a consequence of the differences in educational distributions: In 1970, Germany has many more workers with vocational education than France. Because vocational education links more strongly, Germany's overall linkage is also high. While there are compositional differences, in 1970 the French educational system was not less successful than the German system in providing a close link towards the labor market—it just did so for a smaller share of the labor force. This is the first key result that emerges from the decomposition.

When comparing separately by gender the pattern is similar. A large negative component for the educational margins accounts for a majority of the difference in 1970 for both men and

[6] From this section on, we combine ISCED categories *3ab_voc*, *3c*, and *4* to achieve greater comparability between the two countries.

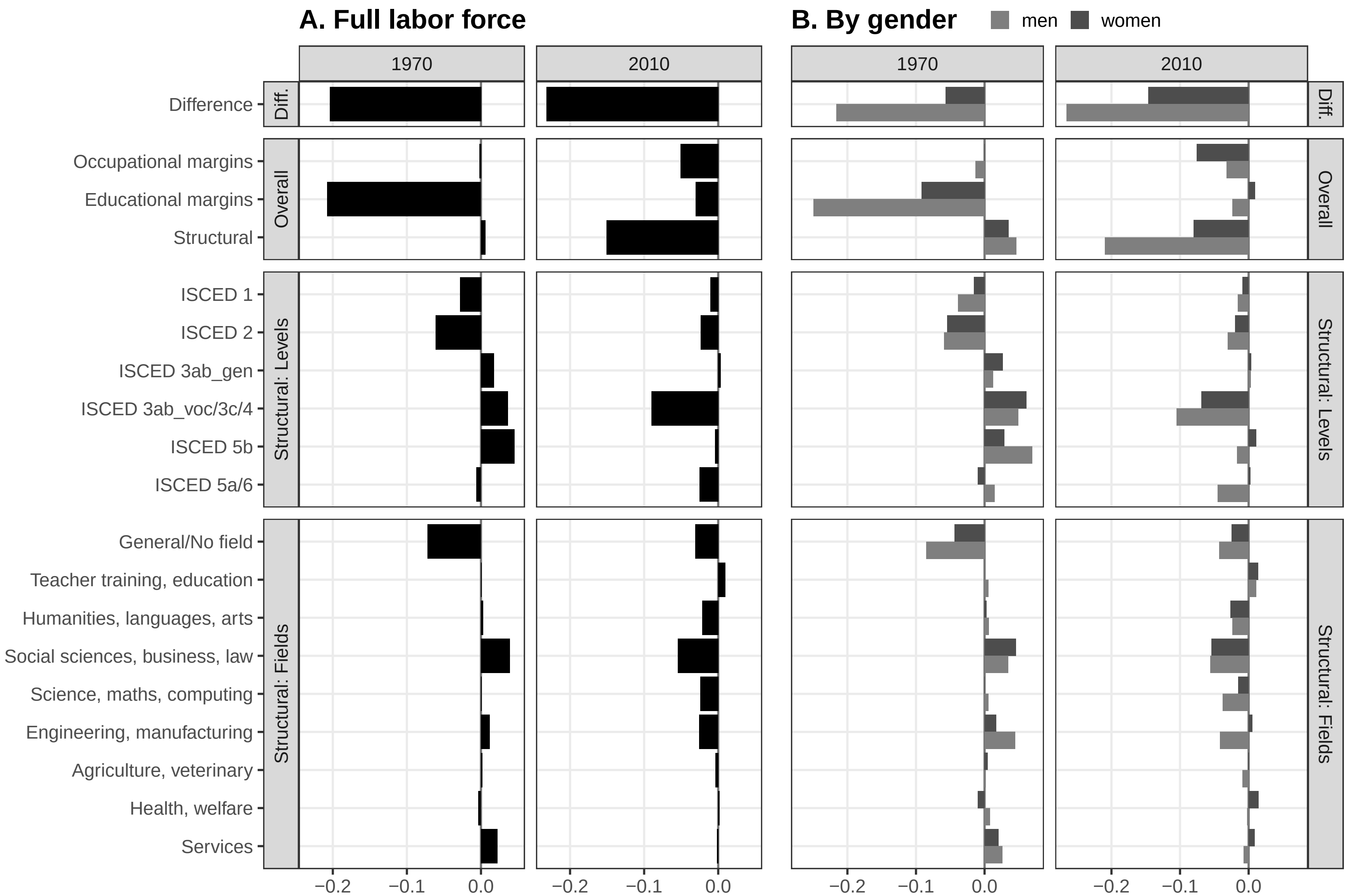

Figure 3.4: Decomposition of M between countries

Note: These results are also available as a table in Appendix A.

women. What the breakdown by gender makes clear is that France had a small advantage in terms of the structural component: Once we account for the compositional differences, the association between education and occupation was *stronger* in France than in Germany. Notably, the cross-national difference in educational margins is much more pronounced for men than for women, which is a consequence of the fact that overall linkage differences are much smaller for women than for men.

In 2010, the pattern has changed markedly: While the country differences in linkage remain on the same order of magnitude, they are now to a much lesser degree explained by compositional differences. Instead, about two thirds of the difference in M between Germany and France are now accounted for by structural differences. Given the convergence in educational distributions between Germany and France since 1970 — mostly reflected by the expansion of vocational training in France, with less change in Germany — it is not surprising that the country differences in the educational marginal distributions contribute much less to the overall country difference in M in 2010 than they did in 1970. As in 2010, the linkage differences are smaller for women than for men.

Given the large size of the structural difference between France and Germany, especially in 2010, it is interesting to study where these differences originate. The two further decompositions reveal a considerable heterogeneity by educational levels. First, the most striking result is that in 1970 graduates of ISCED *3ab_voc/3c/4* and *5b* linked more strongly in France than in Germany; in particular, the strength of the link between school and work was higher for upper-secondary vocational graduates in France than in Germany. This again confirms that, in contrast to what MSS argued, the French vocational system did provide a close link between education and occupation. For tertiary education the differences are smaller. We do find that in 2010 the average graduate with a tertiary degree had higher linkage in Germany compared to France, but the difference is small. The main reason why Germany links much stronger in 2010 than in 1970 stems from upper secondary vocational education (ISCED categories *3ab_voc/3c/4*), which contain both dual training and school-based vocational education. For women, this category accounts for almost all of the difference, while men also gained a small increase in linkage in higher education (ISCED *5a/6*).

The decomposition by field of study further shows that the French advantages in 1970 were concentrated in services, business, engineering, and manufacturing. Especially in engineering and manufacturing, men had higher structural linkage than women, which likely reflects the focus of French vocational training on these stereotypically male occupations. In 2010, higher linkage is visible across most fields in Germany, where we can observe higher structural linkage for men in fields such as science, maths, and computing, as well as engineering and manufacturing. These trends likely reflect the continuing strong occupational segregation by gender in both labor markets.

3.6.4 Within-country change over time

The results from the previous section indicate that there was substantial change over time between the two countries, which implies that at least one of the countries must itself have changed over time. Given that the structural differences between Germany and France are much more pronounced in 2010 than 1970, it is possible that Germany has increased its structural linkage, France has lost some of its structural linkage, or a combination of these two processes. To answer this question, we again use eq. (3.3), but now apply this decomposition to study change within each country over time.

As seen in Figure 3.1, observed total linkage has increased in both countries at roughly comparable rates. However, the decomposition reveals that the reasons for this increase are starkly different. In Germany, the counterfactual scenario shows that the *structure* of linkage between educational programs and occupational destinations is almost equally strong in 1970 and 2010. The rise in linkage strength is due to rapid growth in educational credentials that link more strongly to the occupational structure. To be precise, 86% of the total difference can be explained by changes in the educational and occupational distributions, with the remaining 14% explained by *increasing* structural linkage. In both countries, the occupational marginal com-

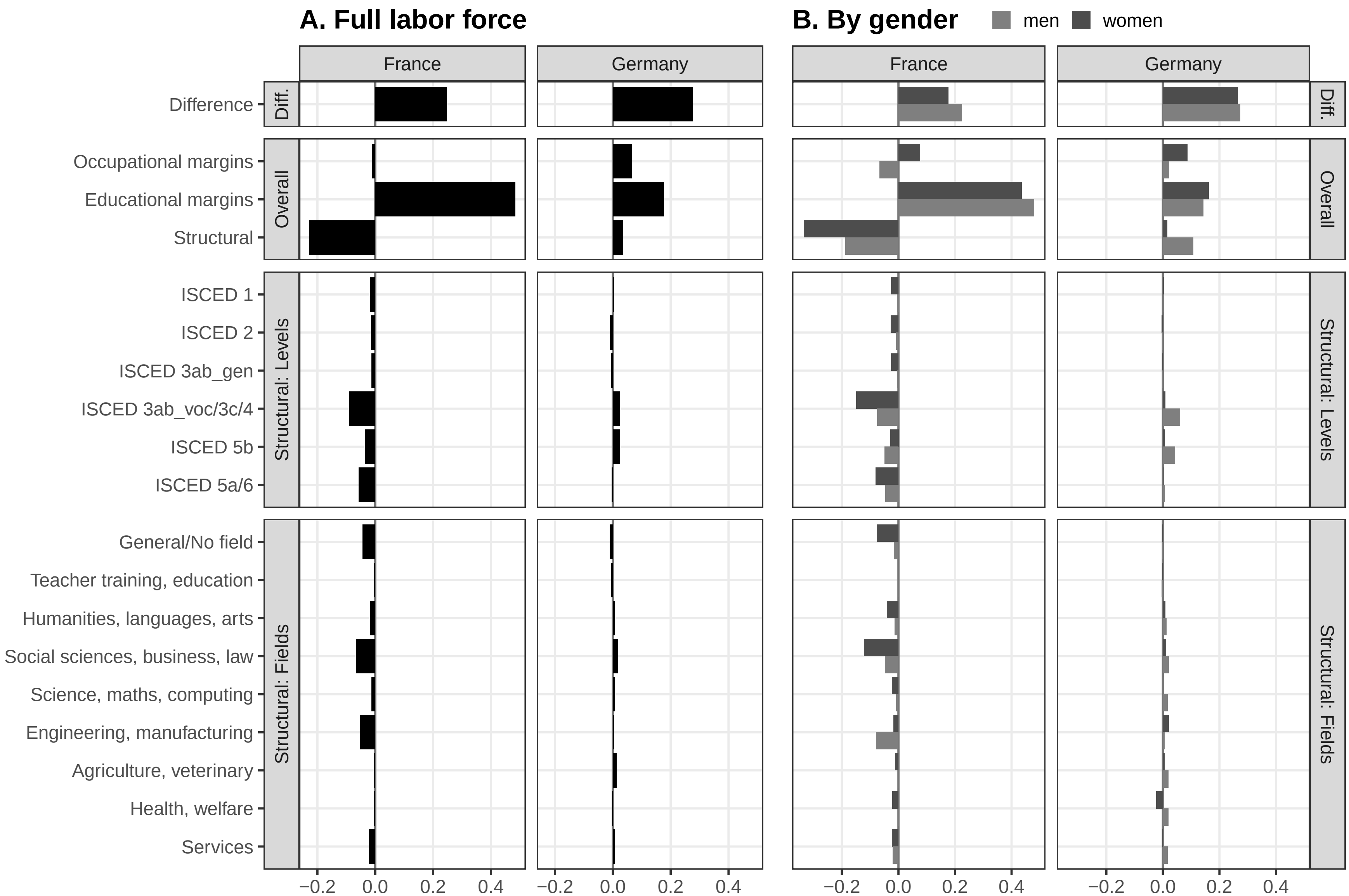

Figure 3.5: Decomposition of M within countries over time

Note: These results are also available as a table in Appendix A.

ponents are higher for women than for men, which is consistent with the idea of women "catching up" to the occupational attainment of men.

In France, marginal and structural change have different signs, and thus partially offset each other. Due to educational expansion, the change in the marginal educational and occupational distributions point to a large *increase* in linkage. In fact, if in France only the educational distribution had changed (without influencing structural linkage), France in 2010 would have almost the same total linkage strength as Germany in 2010. However, a large and *negative* component of structural change offsets a large part of the increase that would be expected from changes in the marginals alone.

Much of the decrease in structural change is due to declines in vocational education (*3ab_voc/3c/4*), and women have been much more affected by declines in structural linkage than men. We again also find that most of the change was concentrated in the fields of business, engineering, and manufacturing. In Germany, the small increase in structural linkage is almost entirely due to the male labor force, and especially concentrated in vocational education. If stronger linkage provides benefits in the labor market, both countries have changed in ways that have been unfavorable for women.

Changes in linkage for tertiary education have been small. This is another strong indication that tertiary education is relatively unaffected from developments in the vocational skill formation system. Both over time and between countries, structural differences in tertiary education are small. In this sense, school-to-work linkages in tertiary education are very similar between these two countries and over time, while the differences for vocational education are more pronounced.

3.7 Discussion: A historical perspective

Using data that span the 40 years from 1970 to 2010, we find that an in-depth comparison of the French and German skill formation systems leads to a more complex picture than commonly assumed. The French skill-formation system went through dramatic change in the

40 years between 1970 and 2010. In 1970 France, only a small part of the workforce obtained vocational education, but this training was effective in providing linkage towards specific occupations—even more so than in Germany at the time! This runs counter to the claims made by MSS. Since then, vocational education has rapidly expanded in France, but this training no longer provides such effective linkage. The country differences have reversed: While vocational training in 1970 France was considerably more effective in providing linkage than Germany's vocational training, vocational education in Germany today is more effective than in France. Our findings highlight the importance of distinguishing between margins and structure: In 1970, Germany was much more vocationally oriented than France, but did not provide higher structural linkage. In 2010, the educational distributions of the two countries have converged, but Germany provides stronger structural linkage.

The literature on educational institutions supports these findings and provides an explanation for the decrease in structural linkage in France. It is generally argued that over the past 50 years, France underwent more extensive educational reforms than Germany. The French expansion both at the vocational and the tertiary level was produced by a series of significant reforms—often with explicit reference to the situation in Germany. Charles Day's (2001) history of French vocational and technical education focuses on the institutional fragmentation of general and technical education. Day traces this back to as early as 1890, when the new technical schools were placed under the direction of the ministry of commerce, instead of the ministry of education. This institutional split deepened, and by 1900, the "technical division" had developed a complete educational system on its own, with its own primary, secondary, and tertiary schools. The technical division catered to a selected group of students and aimed to provide high-level vocational training and schooling—a vast difference to Germany's vocational training that, even in the 1970s, can already be characterized as a system "for the masses." The selective nature of vocationally-trained workers in 1970 France aligns with the high amount of structural linkage that we observe for France in 1970, especially for well-trained vocational graduates.

The situation in France is exemplary of the contentious debate about the goals of general versus skilled education, neatly summarized by the opposing poles of the ministries of commerce, favoring specialized education with the backing of industry, and the ministry of education, arguing for a civic ideal of education. The integration of the two systems, while proposed several times in the course of the 20th century, did not occur until 1960. Since then, various reforms followed that expanded vocational education at different educational levels. Of special note is the "loi Jospin" from 1989, named after the education minister, which formulated the goal that "all young people reach a recognized level of training" (n.a. 1996, p. 49).

Meanwhile, France continued to lack the "decentralized cooperation" between unions, employers, and the state that characterized Germany (Hall and Soskice 2001). It could well be that this lack of decentralized cooperation created imbalances in the supply and demand of workers who were graduating from the expanding specialized educational programs in France. The weaker regulation might also have given French employers greater flexibility in hiring workers whose credentials were not the natural match to the jobs but who otherwise were judged by the employers to be good fits. Either or both of these mechanisms could explain the combination of growing composition-based linkage strength and declining structural-based linkage strength in France.

The contentious history of educational reforms in France contrasts markedly with the German situation (Brauns et al. 1999). The basic system in Germany has been in place since the late 19th century, and despite the two world wars and the political upheavals, has changed only little (Thelen 2004). The German arrangement of vocational ('dual') training has existed since the late 19th century, and relies on close coordination between industry and the educational system. Reforms were planned in Germany as well, especially by the social-democratic governments (1969-1982), and were aimed to remove early tracking or replace some of the vocational training with more general, school-based training. However, these reforms were not implemented. This is clearly visible in our results, which show that in Germany almost no structural change can be observed.

3.8 Conclusions

We investigated how educational systems match workers to the labor market and how this has changed across two countries that are often seen as ideal-types: France and Germany. We summarize our findings in three main points.

First, school-to-work linkages have increased over time in both France and Germany. This means that occupations have become more educationally homogeneous, and educational pathways typically lead to a smaller set of occupations. In Germany, the increase in linkage strength was achieved through a relatively modest quantitative expansion in both vocational and tertiary education, but this expansion was not paired with substantial reforms that would have altered school-to-work linkages. In France on the other hand, we find a large increase in the number of vocational and tertiary graduates. This increase was accompanied by a decline in structural linkage. The decline was especially pronounced for vocational graduates, and was greater for women than for men in both countries. While more students in France enrolled in vocational programs, this came at the cost of the linkage strength of these programs, as students increasingly dispersed across different occupations.

Our results underscore the importance of attending to structural as well as compositional differences in conceptualizing educational systems across countries and over time. This distinction is also important for policymakers and the debate about general vs. vocational education, more generally. The "success" of vocational programs is often defined by a strong linkage to the labor market (indicated, for instance, by a relative ease in entering the labor market). However, faced with changing labor markets and skill requirements, some have called into question the benefits of strong linkage (Hanushek et al. 2017). Because the linkage approach shifts the focus from national education systems to individual educational pathways, it can contribute to this debate by studying the effects of weakly and strongly linking programs on workers' career outcomes (Forster et al. 2016).

Second, we find that the German vocational system in 1970 was not—on average—substantially more efficient in allocating graduates to specific occupations than the French skill formation system. This finding is a major departure from the results presented by MSS that have been reinforced by other studies. The main reason that Germany was portrayed as a vocational system is because more students were enrolled in secondary-level vocational programs in Germany than in France. While there are clear differences in the *rates* of vocational education, the two countries were similar in the extent to which particular educational programs created a strong link to the labor market.

In particular, we find that some of MSS's conclusions are not warranted when the whole of the labor force, including women, is taken into account. We thus emphasize that the historical study of skill-formation systems needs to pay attention to the role of gender. Country differences are more pronounced when only men are considered. For female workers, France and Germany were already very similar in the 1970s and have remained similar since then. This may reflect the common experiences of women in both France and Germany, slowly gaining access to the labor market and higher education. Vocational policy is often especially focused on manufacturing and business degrees, which were historically geared towards men. The fact that many classifications of skill-formation systems are based only on the male workforce affects how we think about skill formation systems. We show that ignoring the female workforce has large consequences for today's conception of skill formation systems, particularly because a large share of educational expansion is caused by an increase in female enrollment in (higher) education.

Our results raise the question whether the cross-national classifications of skill formation systems that are dominant in the current literature do justice to actual cross-national differences. We believe this not to be the case. When looking more closely into how school-to-work linkages are established, countries might be similar on some aspects (structural linkage), but differ on others (composition of workers across the programs). Moreover, the differences *within* countries are as large or larger than differences *between* countries. The characterization of skill-formation systems as qualificational or organizational might thus be even less appropriate when a large fraction of the labor force has attained tertiary education. Future research should move beyond treating countries as entities with homogeneous skill formation systems that are stable over time.

3.9 Appendix

A. Full decomposition

Table 3.3: Decomposition of M between countries

	All workers				Male				Female			
	1970	(%)	2010	(%)	1970	(%)	2010	(%)	1970	(%)	2010	(%)
M Germany	0.64		0.92		0.59		0.87		0.56		0.82	
M France	0.44		0.68		0.38		0.60		0.50		0.68	
Difference	−0.20	(100)	−0.23	(100)	−0.22	(100)	−0.27	(100)	−0.06	(100)	−0.15	(100)
Total marginal	−0.21	(103)	−0.08	(35)	−0.26	(121)	−0.06	(21)	−0.09	(161)	−0.07	(45)
Occupational margins	−0.00	(1)	−0.05	(22)	−0.01	(6)	−0.03	(12)	0.00	(−0)	−0.08	(52)
Educational margins	−0.21	(102)	−0.03	(13)	−0.25	(115)	−0.02	(9)	−0.09	(162)	0.01	(−6)
Total structural	0.01	(−3)	−0.15	(65)	0.05	(−21)	−0.21	(79)	0.03	(−61)	−0.08	(55)
by levels												
ISCED 1	−0.03	(14)	−0.01	(5)	−0.04	(18)	−0.02	(6)	−0.02	(28)	−0.01	(6)
ISCED 2	−0.06	(30)	−0.02	(10)	−0.06	(27)	−0.03	(11)	−0.05	(96)	−0.02	(13)
ISCED 3ab_gen	0.02	(−9)	0.00	(−2)	0.01	(−6)	0.00	(−1)	0.03	(−47)	0.00	(−2)
ISCED 3ab_voc/3c/4	0.04	(−18)	−0.09	(39)	0.05	(−23)	−0.10	(40)	0.06	(−108)	−0.07	(47)
ISCED 5b	0.05	(−22)	−0.00	(2)	0.07	(−32)	−0.02	(6)	0.03	(−51)	0.01	(−8)
ISCED 5a/6	−0.01	(3)	−0.03	(11)	0.01	(−7)	−0.05	(17)	−0.01	(18)	0.00	(−2)
by fields												
General/No field	−0.07	(35)	−0.03	(13)	−0.09	(39)	−0.04	(16)	−0.04	(77)	−0.02	(17)
Teacher training, education	0.00	(−0)	0.01	(−4)	0.01	(−3)	0.01	(−4)			0.01	(−9)
Humanities, languages, arts	0.00	(−1)	−0.02	(9)	0.01	(−3)	−0.02	(9)	0.00	(−5)	−0.03	(18)
Social sciences, business, law	0.04	(−19)	−0.05	(24)	0.03	(−16)	−0.06	(21)	0.05	(−80)	−0.05	(37)
Science, maths, computing	0.00	(−0)	−0.02	(10)	0.01	(−3)	−0.04	(14)	0.00	(−0)	−0.02	(10)
Engineering, manufacturing	0.01	(−6)	−0.03	(11)	0.04	(−21)	−0.04	(16)	0.02	(−29)	0.01	(−4)
Agriculture, veterinary	0.00	(−1)	−0.00	(2)	0.00	(−1)	−0.01	(3)	0.00	(−8)	−0.00	(0)
Health, welfare	−0.00	(2)	0.00	(−1)	0.01	(−4)	−0.00	(1)	−0.01	(18)	0.01	(−10)
Services	0.02	(−11)	−0.00	(1)	0.03	(−12)	−0.01	(3)	0.02	(−36)	0.01	(−6)

Table 3.4: Decomposition of M within countries over time

	All workers				Male				Female			
	Germany	(%)	France	(%)	Germany	(%)	France	(%)	Germany	(%)	France	(%)
M 1970	0.64		0.44		0.59		0.38		0.56		0.50	
M 2010	0.92		0.68		0.87		0.60		0.82		0.68	
Difference	0.28	(100)	0.25	(100)	0.27	(100)	0.23	(100)	0.27	(100)	0.18	(100)
Total marginal	0.24	(88)	0.47	(192)	0.17	(61)	0.41	(183)	0.25	(94)	0.51	(290)
Occupational margins	0.07	(24)	−0.01	(−4)	0.02	(8)	−0.07	(−30)	0.09	(33)	0.08	(43)
Educational margins	0.18	(64)	0.48	(196)	0.14	(52)	0.48	(213)	0.16	(61)	0.44	(247)
Total structural	0.03	(12)	−0.23	(−92)	0.11	(39)	−0.19	(−83)	0.02	(6)	−0.33	(−190)
by levels												
ISCED 1	0.00	(1)	−0.02	(−7)	0.00	(1)	−0.01	(−2)	0.00	(1)	−0.02	(−14)
ISCED 2	−0.01	(−3)	−0.01	(−6)	−0.00	(−1)	−0.01	(−3)	−0.00	(−2)	−0.03	(−15)
ISCED 3ab_gen	−0.00	(−2)	−0.01	(−5)	−0.00	(−1)	−0.00	(−2)	−0.00	(−1)	−0.03	(−14)
ISCED 3ab_voc/3c/4	0.02	(9)	−0.09	(−37)	0.06	(22)	−0.08	(−34)	0.01	(3)	−0.15	(−84)
ISCED 5b	0.02	(9)	−0.04	(−14)	0.04	(16)	−0.05	(−22)	0.01	(3)	−0.03	(−16)
ISCED 5a/6	−0.00	(−1)	−0.06	(−23)	0.01	(3)	−0.05	(−21)	0.00	(1)	−0.08	(−45)
by fields												
General/No field	−0.01	(−4)	−0.04	(−18)	−0.00	(−1)	−0.02	(−7)	−0.00	(−1)	−0.08	(−44)
Teacher training, education	−0.01	(−2)	−0.00	(−1)	−0.00	(−2)	−0.00	(−1)	−0.00	(−1)		
Humanities, languages, arts	0.01	(2)	−0.02	(−7)	0.01	(5)	−0.01	(−6)	0.01	(3)	−0.04	(−23)
Social sciences, business, law	0.02	(6)	−0.07	(−27)	0.02	(8)	−0.05	(−21)	0.01	(4)	−0.12	(−69)
Science, maths, computing	0.01	(2)	−0.01	(−5)	0.02	(6)	−0.01	(−3)	−0.00	(−0)	−0.02	(−13)
Engineering, manufacturing	0.00	(1)	−0.05	(−21)	0.01	(2)	−0.08	(−35)	0.02	(8)	−0.02	(−10)
Agriculture, veterinary	0.01	(5)	−0.00	(−2)	0.02	(7)	−0.00	(−1)	0.01	(2)	−0.01	(−7)
Health, welfare	−0.00	(−1)	−0.00	(−2)	0.02	(7)	0.00	(0)	−0.02	(−9)	−0.02	(−12)
Services	0.01	(2)	−0.02	(−9)	0.02	(6)	−0.02	(−9)	−0.00	(−1)	−0.02	(−13)

B. Crosswalks for fields of study

Overview

Survey	**Field of study**	**Occupation**
France		
FQP, 1970	different schemes for vocational education (47 categories) and higher education (10 categories)	manually coded into ISCO-88 using the "profession" variable, as well as the socio-professional categories and the sector of work
FQP, 1985	different schemes for vocational education (47 categories, same as 1970) and higher education (9 categories)	coded as PCS-1982 (*Professions et catégories socioprofessionnelles*), coded into ISCO-88 using a proportional crosswalk obtained from the Enquête Emploi data
Enquête Emploi, 1995-2001	coded as NSF (Nomenclature de spécialités de formation)	coded as PCS-1982, and additionally as ISCO-88 in 1998-2002; the double-coded years were used to create a proportional crosswalk that was applied to the years 1995-1997 and to the 1985 data
Germany		
Census, 1970	different schemes for job of dual training (2-digit KldB [*Klassifizierung der Berufe*] 1970), and vocational schools/higher education (95 categories)	coded as KldB 1970 (very similar to KldB 1988), then coded into ISCO-88 using a proportional crosswalk obtained from the 2006 BIBB/BAuA survey
Census, 1987	different schemes for job of dual training (2-digit KldB 1975), and vocational schools/higher education (similar to 1970)	coded as KldB 1975 (very similar to KldB 1988), then coded into ISCO-88 using a proportional crosswalk obtained from the 2006 BIBB/BAuA survey

France: Nomenclature de spécialités de formation (NSF)

ISCED	NSF
Teacher training and education science	P4
Humanities, languages and arts	C6-C8; D1-D7
Social sciences, business and law	C1-C5; C9; M1-M6; N1-N7
Science, mathematics and computing	B1-B9
Engineering, manufacturing and construction	E1-E2; G1-G8; H1-H5; J1-J4; K1-K6
Agriculture and veterinary	F1-F5
Health and welfare	P1-P3
Services	L1; P5-P7; Q1-Q7; R1-R6; S1-S3

France: 1970/1985 schemes for university degrees

ISCED	1970 scheme	1985 scheme
Teacher training and education science	–	–
Humanities, languages and arts	1 Littéraire : lettres classiques ou modernes, langues, philosophie, psychologie, sociologie, histoire, géographie 8 Artistique	1 Lettres, sciences humaines : lettres classiques ou modernes, langues, philosophie, psychologie, sociologie, histoire, géographie, etc. 8 Architecture, arts, arts appliqués, esthétique industrielle
Social sciences, business and law	2 Juridique, économique, commerciale	2 Secrétariat, documentation, information, journalisme 3 Droit, notariat, carrières administratives et juridiques 4 Sciences économiques, gestion, commerce, comptabilité
Science, mathematics and computing	3 Scientifique I : mathématiques, physique, chimie 4 Scientifique II : sciences naturelles, biologie, zoologie, botanique, géologie	5 Sciences exactes : mathématiques, physique, chimie, informatique, électricité, électronique, optique, mécanique, génie civil et autres sciences de l'ingénieur 6 Sciences de la nature et de la vie : géologie, sciences naturelles, biochimie, biologie, géologie, zoologie, agronomie
Engineering, manufacturing and construction	–	–
Agriculture and veterinary	–	–
Health and welfare	5 Médicale : médecine, pharmacie, dentaire, vétérinaire 6 Paramédicale ou sociale (assistantes sociales, kinésithérapeutes, éducateurs spécialisés, infirmières, sages-femmes, etc.)	7 Médecine, pharmacie, dentaire, vétérinaire (paramedical/social degrees can be inferred from DIP variable)
Services	0 Divers	9 Autres (éducation physique et sportive, enseignement ménager familial, théologie, militaire à l'exclusion des formations des écoles de santé de l'armée classées en 7, etc.)

France: 1970/1985 scheme for non-university degrees

ISCED	1970/1985, non-university degrees
Teacher training and education science	35 Enseignement, animation à caractère éducatif; 42 Formations aux fonctions d'encadrement
Humanities, languages and arts	36 Arts et arts appliqués, esthétique industrielle; 43 Formations littéraires et linguistique
Social sciences, business and law	28 Organisation du travail, gestion et contrôle de la production; 29 Techniques administratives et juridiques appliquée; 30 Secrétariat, dactylographie, sténographie; 31 Techniques financières ou comptables, mécanographie-comptable; 32 Traitement électromécanique et électronique de l'information; 33 Commerce et distribution; 34 Information, documentation, relations publique; 44 Formations économiques, commerciales, juridiques générales ou en gestion des collectivités publiques ou des entreprise; 14 Photographie, industries graphiques (photogravure, composition, impression)
Science, mathematics and computing	16 Chimie, physique, biochimie, biologie, production chimique; 45 Formations générales en sciences ou techniques industrielles
Engineering, manufacturing and construction	03 Mines et carrières (extraction), travail des pierres; 04 Génie civil, travaux publics, topographie; 05 Construction en bâtiment; 06 Couverture, plomberie, chauffage; 07 Peinture en bâtiment, peinture industrielle; 08 Production et première transformation des métaux; 09 Forge, chaudronnerie, constructions métalliques, formations connexes; 10 Mécanique générale et de précision, travail sur machines-outils, automatismes; 11 Électricité, électrotechnique, électromécanique; 12 Électronique; 13 Verre et céramique; 15 Papier et carton; 17 Boulangerie, pâtisserie; 18 Abattage, travail des viandes; 19 Autres spécialités de l'alimentation; 20 Textiles; 21 Habillement, travail des étoffes; 22 Travail des cuirs et peaux; 23 Travail du bois; 24 Conducteurs d'engins terrestres : engins de chantiers, de levage, de transport et machines agricoles; 25 Autres formations des secteurs primaires et secondaires (conducteurs de fours, de chaudières, manutention, etc.); 26 Dessinateurs du bâtiment et des travaux publics; 27 Dessinateurs industriels
Agriculture and veterinary	01 Agriculture, élevage, forestage 02 Pêche, navigation maritime et fluviale
Health and welfare	37 Santé, secteur para-médical, services sociaux
Services	38 Soins personnels; 39 Services dans l'hôtellerie et les collectivité; 40 Arts ménagers (Hauswirtschaft); 41 Surveillance, sécurité; 46 Préformations, formations générales à finalité professionnelle

Germany: 1970/1987 schemes for university degrees

ISCED	1970 scheme	1987 scheme
Teacher training and education science	61-69 Erziehung, Lehramt; 70-72 Theologie	64-71 Erziehung, Lehramt; 72-74 Theologie
Humanities, languages and arts	76 Philosophie; 77 Psychologie; 78 Geschichte, Völkerkunde; 79 Bibliothek/Archiv/Publizistik; 80-90 Philologie und Sprachen; 91-94 Kunst, Musik, etc.	75 Philosophie; 76 Geschichte; 77 Bibliothek/Publizistik; 78-85 Sprachen; 86 Psychologie; 87-91, Kunst, Musik, etc.
Social sciences, business and law	31 Warenhandel; 32 Bank/Versicherung; 34-35 Kaufm. Handelsschulen; 41 Verwaltung und Organisation; 42 Rechtwesen; 44-50 Wirtschafts- und Sozialwissenschaften	34 Unternehmensf.; 35 Kaufm. Verwaltung; 36 Warenhandel; 37 Bank/Versicherung; 40 Öffentliche Verw.; 42 Rechtswesen 43-46 Wirtschafts- und Sozialwissenschaften
Science, mathematics and computing	25 Mathematik und math. Technik; 26-28 Naturwissenschaften; 29-30 Wirtschafts- und Betriebstechnik; 24 Sonstige industrielle/handwerkliche Fachrichtungen	26 Mathematik und math. Technik; 27 DV 28-30 Naturwissenschaften; 31-32 Wirtschafts- und Betriebstechnik
Engineering, manufacturing and construction	4 Nahrungsmittel; 5 Berg/Hüttenwesen; 6 Stein Keramik Glas; 7-9 Hoch- und Tiefbau; 10 Bauausstattung; 11 Vermessung; 12 Werkstoffbearbeitung; 13 Feinmechanik, Optik; 14 Maschinenbau; 15 Fahrzeugbau; 16 Elektrotechnik; 17 Holzverarbeitung; 18 Papierherstellung; 19 Druck/Photographie; 20 Graphik; 21 Textil, Leder	4-5 Nahrungmittel; 7 Berg/Hüttenwesen; 8 Stein, Keramik, Glas; 9-11 Hoch- und Tiefbau; 12-13 Bauwesen; 14 Metallbearbeitung; 15 Kunststoff; 16 Feinmechanik, Optik; 17 Maschinenbau; 18 Fahrzeugbau; 19 Elektro; 20 Holzverarbeitung; 21 Papierherstellung; 22 Druck; 23 Techn. Zeichnen; 24 Textil, Leder
Agriculture and veterinary	1 Land- und Tierwirtschaft; 2 Garten- und Weinbau; 3 Forst/Fischerei	1 Land- und Tierwirtschaft; 2 Garten- und Weinbau; 03 Forst/Fischerei
Health and welfare	51-54 Medizin und Pharmazie; 55 Krankenfürsorge; 56-60 Sozialarbeit, Sozialpflege	53-56 Medizin und Pharmazie; 57 Krankenpflege; 58 MTA; 59 nichtärtzliche Heilbehandlung; 60-62 Sozialarbeit, Sozialpflege
Services	36 Gaststätten; 37 Hauswirtschaft; 38 Bau- und Strassenreinigung; 39+40 Körperpflege; 33 Verkehr; 43 Ordnungs- und Sicherheitswahrer	48 Hotel/Gaststätten.; 49 Touristik; 50 Reinigung; 51-41 Körperpflege; 38 Verkehr; 41 Öff. Sicherheit und Ordnung

Germany: 1970/1987 scheme for non-university degrees (based on KldB 1970/1975)

ISCED	1970/1987, non-university degrees
Teacher training and education science	87 Lehrer
Humanities, languages and arts	82 Publizisten, Dolmetscher, Bibliothekare; 83 Künstler und zugeordnete Berufe
Social sciences, business and law	68 Warenkaufleute; 69 Bank-, Versicherungskaufleute; 70 Andere Dienstleistungskaufleute; 75 Unternehmer, Wirtschaftsprf.l 76 Abgeordnete, administrativ entscheidende Berufstätige; 77 Rechnungskaufleute, DV; 78 Bürofach-, Bürohilfskräfte; 81 Rechtsverwahrer, -berater; 88 Geistes- und naturwissenschaftliche Berufe
Science, mathematics and computing	61 Chemiker, Chemieingenieure/Physiker, Physikingenieure, Mathematiker
Engineering, manufacturing and construction	7-9 Bergleute, Mineral, Erdöl; 10-13 Stein, Keramik, Glas; 14 Chemiearbeiter; 15 Kunststoffverarbeiter; 16 Papierhersteller, -verarbeiter; 17 Drucker; 18 Holzaufbereiter; 60 Ingenieure; 62 Techniker; 63 technische Sonderkräfte; 19-26 Metall; 27 Schlosser; 28 Mechaniker; 29 Werkzeugmacher; 30 Metallfeinbauer etc.; 31 Elektriker; 32 Montierer und Metallberufe, a.n.g.; 33 Spinnberufe; 34-36 Textil; 37 Lederhersteller, Felle; 39 Back-Konditorwarenhersteller; 40 Fleisch-, Fischverarbeiter; 41 Speisenbereiter; 42 Getränke-, Genußmittelhersteller; 43 Übrige Ernährungsberufe; 44-51 Bauberufe; 52 Warenprüfer, Versandfertigmacher; 53 Hilfsarbeiter; 54 Maschinisten etc.
Agriculture and veterinary	1 Landwirte; 2 Tierzüchter, Fischereiberufe; 3 Verwalter, Berater in der Landwirtschaft und Tierzucht; 4 Landwirtschaftliche Arbeitskräfte, Tierpfleger; 5 Gartenbauer; 6 Forst- Jagdberufe
Health and welfare	84 Ärzte, Apotheker; 85 Übrige Gesundheitsdienstb.; 86 Sozialpflegerische Berufe; 89 Seelsorger
Services	90 Körperpfleger; 91 Gästebetreuer; 92 Hauswirtschaftliche Berufe; 93 Reinigungsberufe; 71-73 Verkehr; 74 Lagerverwalter, Lager-, Transportarbeiter; 79 Dienst-, Wachberufe 80 Sicherheitsverwahrer

C. Estimates of uncertainty for proportional crosswalks

To harmonize occupational codings, we use proportional crosswalks for the German datasets for 1970 and 1987, as well as for the French dataset for 1985, and for the early years of the French Labor Force Survey (1995-1997). As explained in the main text, these crosswalks add a small amount of uncertainty to our estimates in these years. To estimate these uncertainties, we repeat the matching process 100 times. This creates 100 equally plausible occupational

Table 3.5: Means and standard deviations of local linkage scores across 100 iterations

		Germany, 1970		Germany, 1987		France, 1985		France, 1996	
ISCED	Field	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ISCED 1	General/No field	1.026	0.006	0.631	0.003	0.354	0.001	0.398	0.000
ISCED 2	General/No field	0.243	0.000	0.280	0.000	0.072	0.000	0.135	0.000
ISCED 3ab_voc	Humanities, languages, arts	1.190	0.008	0.987	0.009	2.058	0.000	1.625	0.016
ISCED 3ab_voc	Social sciences, business, law	0.701	0.000	0.634	0.000	0.945	0.004	0.608	0.001
ISCED 3ab_voc	Science, maths, computing	1.012	0.005	0.812	0.008	2.135	0.002	1.642	0.027
ISCED 3ab_voc	Engineering, manufacturing	0.384	0.000	0.389	0.000	1.080	0.004	0.611	0.002
ISCED 3ab_voc	Agriculture, veterinary	1.063	0.002	1.631	0.003	1.450	0.008	1.358	0.006
ISCED 3ab_voc	Health, welfare	1.340	0.003	1.727	0.001	1.385	0.001	1.415	0.013
ISCED 3ab_voc	Services	0.864	0.001	0.820	0.001	2.024	0.084	1.559	0.008
ISCED 3ab_gen	General/No field	0.800	0.002	0.502	0.001	0.644	0.001	0.369	0.001
ISCED 3c/4	Humanities, languages, arts	3.433	0.040	2.366	0.051	1.636	0.043	1.018	0.015
ISCED 3c/4	Social sciences, business, law	1.423	0.003	0.972	0.002	0.735	0.001	0.455	0.000
ISCED 3c/4	Science, maths, computing	4.824	0.377			1.912	0.021	1.871	0.053
ISCED 3c/4	Engineering, manufacturing	0.599	0.008	0.531	0.002	0.385	0.001	0.421	0.000
ISCED 3c/4	Agriculture, veterinary	1.256	0.028	2.124	0.028	1.083	0.003	1.312	0.001
ISCED 3c/4	Health, welfare	1.219	0.025	2.306	0.005	1.737	0.005	1.615	0.004
ISCED 3c/4	Services	1.302	0.011	1.097	0.011	1.143	0.002	0.720	0.001
ISCED 5b	Humanities, languages, arts	0.913	0.010	0.675	0.015	1.206	0.005	1.034	0.003
ISCED 5b	Social sciences, business, law	0.656	0.001	0.616	0.001	1.070	0.005	0.684	0.001
ISCED 5b	Science, maths, computing	0.971	0.005	0.948	0.004	1.216	0.005	1.066	0.010
ISCED 5b	Engineering, manufacturing	0.436	0.001	0.460	0.001	4.649	0.000	0.985	0.002
ISCED 5b	Agriculture, veterinary	0.955	0.006	2.074	0.004			1.247	0.010
ISCED 5b	Health, welfare	2.316	0.002	2.110	0.002	2.276	0.001	2.301	0.001
ISCED 5b	Services	0.792	0.003	1.236	0.003	1.778	0.009	0.832	0.009
ISCED 5a/6	Teacher training, education	3.051	0.003	2.428	0.002			2.695	0.022
ISCED 5a/6	Humanities, languages, arts	2.422	0.007	1.694	0.005	1.953	0.001	1.678	0.001
ISCED 5a/6	Social sciences, business, law	2.154	0.003	1.203	0.001	1.868	0.002	1.196	0.001
ISCED 5a/6	Science, maths, computing	2.407	0.006	1.562	0.003	2.103	0.001	1.764	0.001
ISCED 5a/6	Engineering, manufacturing	2.189	0.002	1.490	0.002			1.848	0.002
ISCED 5a/6	Agriculture, veterinary	1.783	0.012	1.583	0.008			1.955	0.004
ISCED 5a/6	Health, welfare	4.016	0.002	2.993	0.002	4.116	0.000	3.679	0.000
ISCED 5a/6	Services	1.636	0.015	1.564	0.004	2.486	0.011	1.851	0.005

distributions for the four country-years. For each country-year and iteration, we then calculate the M and local linkage measures. Table 3.5 reports the means and standard deviations of the local linkage scores across the iterations. The standard deviations of the M values (not shown) are all below 0.001. Given the small standard errors of our estimates, we proceed by reporting the values of one iteration only in the main text.

D. Lower age cutoff

Figure 3.6 shows the effects that different age cutoffs have on the overall M. As the differences are minor, we continue to use the cutoff at age 15.

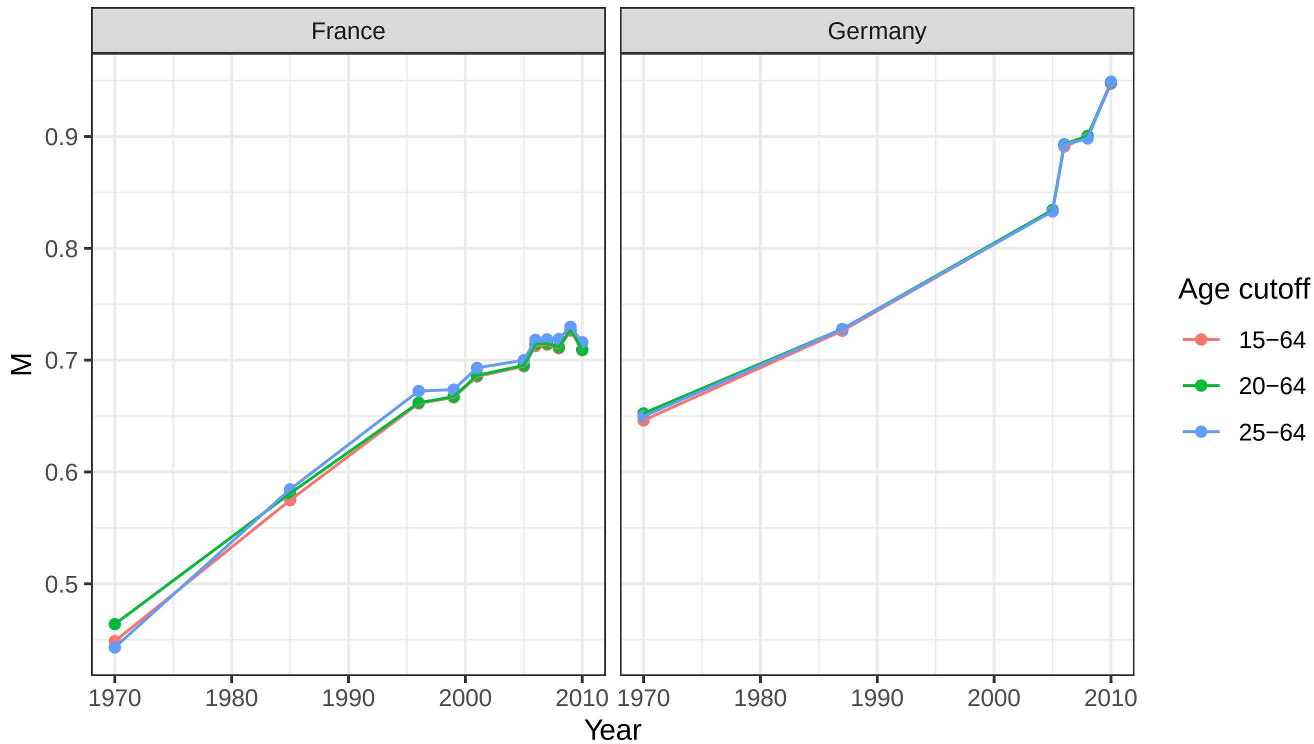

Figure 3.6: M values for different age cutoffs

E. Germany, 2005-2010

The reader may have noted that the linkage measure increased strongly in Germany between 2005 and 2010. The increase is especially strong between 2005 and 2006.

Figure 1 in the main text may provide a partial answer to the question of why linkage increased so much between 2005 and 2006 in Germany, as the figure shows that the increases have been most pronounced for the youngest age group, 15-34. This increase might thus be a process brought about by cohort change. As we write in the paper, a full exploration of age, period, and cohort effects is beyond the scope of this study, but it shows again that such an undertaking could be very useful.

Another explanation is data noncomparability across certain years. The EU-LFS, on which the data for 2005-2015 is based changed "changed [in 2006] significantly to reduce the burden on respondents; since then, all variables have to be collected on a yearly basis, but only a selection of them on a quarterly basis".[7] Germany specifically changed from an annual sample in 2005 to a quarterly sample in 2006. This is clearly visible in the number of cases, too: In 2005, we have almost 185,000 cases in Germany, while in 2006 we only have 40,000. While it is not obvious why these differences would influence the linkage measure, a change in survey methods can be a cause for some concern.

Table 3.6: Sensitivity analyses for 2005/2010

	Germany		France-Germany	
	1970-2010	1970-2005	2010	2005
M1	0.642	0.642	0.917	0.816
M2	0.917	0.816	0.685	0.675
Difference	0.275	0.175	-0.232	-0.142
Educ. marginals	0.174	0.148	-0.023	-0.028
Occ. marginals	0.064	0.071	-0.041	-0.035
Structural	0.037	-0.042	-0.152	-0.069

Table 3.6 reports a sensitivity analysis for the German data. What if we only had data up until 2005? Would this have changed our results? We first look at the within-country decompositions over time. The first column shows the decomposition that we reported in the paper, comparing Germany in 2010 with Germany in 1970. Our main conclusion here was that the difference is mostly explained by educational expansion, without much structural change. The next column shows the decomposition if we instead choose 2005 as the final year. In some sense, this strengthens our main point, as now a larger part of the difference is explained by the educational margins (.173/.275 ≈ 63% vs .148/.175 ≈ 85%). However, we also note that structural change over this period was negative, instead of positive when studying the whole period.

[7] https://www.gesis.org/en/missy/metadata/EU-LFS/

Apparently, the small increase in structural change that we observe in Germany is a product of recent years, i.e. 2005 to 2010.

We also recalculate the between-country decompositions. Again, column 3 shows the decomposition that we reported in the paper, where “M1” refers to the M for Germany in 2010, and “M2” refers to the M in France in 2010. Here the largest component was structural change, accounting for $.157/.240 \approx 65\%$ of the difference. If we choose 2005 as the comparison year, we find that structural change accounts for $.078/.154 \approx 50\%$. This is slightly smaller, but expected, as we saw that Germany’s structural linkage increase especially between 2005 and 2010.

Overall, we conclude that using 2005 instead of 2010 as the baseline year would not have changed our substantive conclusions. When data for future years become available, it will be possible to assess more definitively whether the apparent increase in structural change in recent years in Germany is real and indeed is being driven by changes in the experience of the most recent entrants to the labor market.

F. Labor force participation

One reviewer raised the question of whether the relatively large increase in linkage strength for older French women could be related to changes in their pattern of labor force participation, particularly as it relates to education. By definition, linkage strength can only be assessed for the employed population. In this sense the results reported in Figure 1 are true regardless of the fraction of French or German adults who are not in the labor force and regardless of their educational distribution. At the same time, it is interesting to observe whether changes in participation patterns differ between Germany and France, particularly for women. Table 3 reports labor force participation rates in the two countries at the beginning and end of our observation window, by gender. Table 3 shows that in both countries the lowest participation rates are for the least educated, and that participation rates among the least educated have fallen, even as this group has become a smaller share of the population. Table 3 also reports generally lower participation rates for women than for men in both countries. The gender gap in participation was considerably larger in Germany than in France in 1970, but this gap has shrunk, and highly educated women are now as likely to participate in the labor market in Germany as in France. In both countries, it would seem plausible that the rise in linkage strength for women would have been slower had more lower-educated women been willing and able to find jobs in the labor force in both countries. At the same time, the upward trends reported in Figure 1 demonstrate strong upward shifts in linkage strength for women given the actual jobs in which they worked and their actual educational distribution.

Table 3.7: Labor force participation rates

	Men				Women			
	Germany		France		Germany		France	
ISCED	1970	2010	1970	2010	1970	2010	1970	2010
5a/6	95	89	92	79	74	80	72	73
5b	93	89	69	85	69	81	59	78
3ab_voc/3c/4	92	80	93	71	51	71	63	61
1/2/3ab_gen	50	33	82	49	32	28	46	40
all	76	70	84	66	41	61	49	58

Chapter 4: A Formal Introduction to the Shapley Decomposition

4.1 Introduction

The decomposition developed and applied in the previous two chapters has the advantage that it works for any comparison between two M indices. The distinction between changes produced by the odds ratios and by the marginal distributions is a long-standing issue in categorical data analysis, and therefore connects to the methodological literature on log-linear modeling. This decomposition can often be a good first step when approaching a segregation problem, especially to rule out that changes in segregation are entirely produced by changes in the marginal distribution(s). However, the interpretation of the terms of the decomposition can be difficult, as they are not connected to the underlying social processes that are usually considered when hypothesizing about the forces that bring about segregation change. In the context of residential segregation, for instance, the marginal-structural decomposition does not tell us whether segregation declined because majority groups moved to minority neighborhoods, minority groups moved to minority neighborhoods, or both. The goal of this chapter is develop a flexible methodology to address questions such as these. The following two chapters then apply this methodology to long-term changes in residential segregation and more recent changes in school segregation in the U.S.

The solution that is considered here is the "Shapley value," named after the game theorist Lloyd Shapley (Shapley 1953). While first developed in the context of game theory to quantify each player's payoff for a game, its applications are far more wide-ranging. Unlike most decomposition procedures that focus on a certain class of problems (such as regression decomposition), the Shapley decomposition defines a general decomposition rule for any given problem. While this chapter is written with segregation decomposition in mind, the Shapley decomposition can be used to decompose many other outcomes that are of interest across the social sciences, such as inequality indices or regression coefficients. In this sense, the Shapley value defines a "meta" decomposition rule, and is is often possible to derive existing procedures, such as Oaxaca-Blinder-Kitagawa regression decomposition (Jann 2008; Kitagawa 1955), from the Shapley value. Given the useful properties of the Shapley value discussed below, a decomposition procedure that cannot be derived from the Shapley decomposition should be subject to heightened scrutiny.

The Shapley decomposition solves two important problems that occur in decomposition analysis. First, it avoids the *path dependency problem*. A classic example to illustrate this problem is the decomposition of the R^2 of a regression model. Let $x_1, \ldots, x_m$ be a set of m predictors that are jointly entered into a regression model to produce some value of the R^2. One question of interest could then be how much each of the predictors contributes to the R^2 of the full model. A simple solution to obtain such a solution is to enter the predictors one-by-one, i.e. we start with a model that just includes x_1, record its R^2, and assign this contribution to the predictor x_1. We then enter another predictor, x_2, again record the R^2, and assign the difference between this R^2 and the R^2 from the previous model as the contribution of predictor x_2. Continuing in this way will yield an additive decomposition of the R^2 of the full model, where, arguably, each contribution depends only on the value of the added predictor. However, unless all predictors are orthogonal to each other (unlikely in any practical problem), the contributions will depend on the order in which the predictors are entered. For instance, instead of entering the predictors in the sequence $(x_1, x_2, \ldots, x_m)$, one could enter the predictors in the sequence $(x_m, x_{m-1}, \ldots, x_1)$, yielding a different set of contributions. In general, any permutation of the sequence will yield different contributions, which defines the path dependency problem (Fortin et al. 2011). The path dependency can sometimes be avoided if there is a theoretical reason to enter the predictors in a certain sequence, but in practice, such situations will be rare.

Another solution to avoid the path dependency problem is to define a base model—this could be the empty model that includes no predictors, or the full model that includes all predictors—and to then include/exclude one predictor at a time. For instance, one would find the contribution of predictor x_i by just entering x_i into the model, and recording the R^2 as the contribution of x_i. Hence, all the contributions would be defined on a simple regression model, just varying the single predictor. This approach has two problems: First, depending on the base model, the contributions of each predictor will differ, and there is usually no compelling reason to prefer one base model over another. Second, this rule does not yield a decomposition in the sense that the sum of the contributions equal the R^2 of the full model. This is called the *efficiency problem*. This problem is sometimes "solved" by introducing a residual term. Interpreting such a residual term is often difficult, and if it is large (either very positive or very negative compared to the contributions), the decomposition will not be informative.

Both of these problems only occur for nonlinear outcomes, such as R^2. For such nonlinear outcomes, the Shapley decomposition provides a unique solution to find the contributions, avoiding both problems.

4.2 Shapley decomposition

To define the Shapley decomposition formally (the notation loosely follows Shorrocks 2013), let I be the outcome of interest, $N = \{1, 2, \ldots, m\}$ be the set of factors of interest ("players" in game theory), and $v(\cdot)$ be a set function, called the value function in game theory, whose inputs jointly determine the outcome. The outcome of interest I can then be written as

$$I = v(N) - v(\varnothing) = v(\{1, 2, \ldots, m\}) - v(\varnothing).$$

In the above example, I is the R^2 of the full model, and $v(\cdot)$ returns the R^2 for a regression model where only the input factors are entered as predictors. For instance, $v(\{1, 3\})$ would return the R^2 for a model that includes only the predictors x_1 and x_3. For the case of R^2, $v(\varnothing) = 0$, as including no predictors yields an R^2 of zero.

The goal of any additive decomposition procedure is to find appropriate values for the contributions $\varphi_1, \varphi_2, \ldots, \varphi_m$ that satisfy

$$I = \varphi_1 + \varphi_2 + ... + \varphi_m. \tag{4.1}$$

A “naive” version that fulfills (4.1) is a decomposition that enters all factors sequentially, but, as discussed above, the contributions will then depend on the order in which the factors are entered. Formally, let $\gamma \subseteq N \setminus \{i\}$, i.e. γ is any subset of the factors not involving i. Then the marginal contribution of factor i towards γ is

$$\delta_i(\gamma) = v(\gamma \cup \{i\}) - v(\gamma).$$

Unless v is a linear function, the $\delta_i(\gamma)$ will differ for different values of γ.

The solution of the Shapley decomposition is to consider *all* the sequences in which the factors could be entered, arriving at the following decomposition rule for a set of players N and a value function v:

$$\begin{aligned}\varphi_i(N,v) &= \frac{1}{m!}\sum_{\gamma_j \subseteq N\setminus\{i\}} |\gamma_j|!(m-1-|\gamma_j|)!\delta_i(\gamma_j) \\ &= \frac{1}{m!}\sum_{\gamma_j \subseteq N\setminus\{i\}} |\gamma_j|!(m-1-|\gamma_j|)!\left[v(\gamma_j \cup \{i\}) - v(\gamma_j)\right] \end{aligned} \tag{4.2}$$

where the summation extends over all possible subsets γ_j of $S \setminus \{i\}$, and $|\cdot|$ denotes the cardinality of a set. The complexity of the formula should not distract from the fact that the idea is very simple. As Shorrocks (2013) writes,

> “In broad terms, the proposed solution considers the marginal effect on I of eliminating each of the contributory factors in sequence, and then assigns to each factor the average of its marginal contributions in all possible elimination sequences. This procedure yields an exact additive decomposition of I into m contributions.” (p. 101)

As an example, consider an outcome that we would like to decompose into two factors:

$$I = v(\{1,2\})$$

At this point, we define two $v(1)$ and $v(2)$, which reflect the situation if only factor 1 (2) is included in the calculation. Then, there are two possible elimination sequences for each factor.

To eliminate factor 1, we can compare $v(\{1,2\})$ to $v(2)$, as well as $v(1)$ to $v(\varnothing)$. To quantify the contribution φ_1 of factor 1, the average over the two possible elimination sequences is taken:

$$\varphi_1 = \tfrac{1}{2}\left[v(\{1,2\}) - v(2)\right] + \tfrac{1}{2}\left[v(1) - v(\varnothing)\right]$$

The elimination of factor 2 proceeds in the same way:

$$\varphi_2 = \tfrac{1}{2}\left[v(\{2,1\}) - v(1)\right] + \tfrac{1}{2}\left[v(2) - v(\varnothing)\right]$$

Because $v(\{1,2\}) = v(\{2,1\})$, it is immediately evident that

$$\varphi_1 + \varphi_2 = v(\{1,2\}) - v(\varnothing).$$

If we define $v(\varnothing) = 0$, then $I = \varphi_1 + \varphi_2$.

4.3 Linear and nonlinear functions

If $v(\cdot)$ is a linear function, then it is easy to verify that (4.2) reduces to the expected result. Define the factors $N = \{1,2,\ldots,m\}$ and a linear function of the form $v(\gamma) = \sum_{i\in\gamma} w_i x_i$, such that factor 1 maps to the contribution of $w_1 x_1$, factor 2 maps to the contribution of $w_2 x_2$, and so on. The weights, w_i, are included for generality. Under these conditions, $v(\gamma_j \cup \{i\}) - v(\gamma_j) = w_i x_i$, and (4.2) can be simplified as follows:[1]

$$\begin{aligned}
\varphi_i &= \frac{1}{m!} \sum_{\gamma_j \subseteq N\setminus\{i\}} |\gamma_j|!(m-1-|\gamma_j|)!\, w_i x_i \\
&= \frac{w_i x_i}{m!} \sum_{l=0}^{m-1} \binom{m-1}{l} l!(m-1-l)! \\
&= \frac{w_i x_i}{m!} \sum_{l=0}^{m-1} (m-1)! \\
&= \frac{w_i x_i}{m!} m(m-1)! \\
&= w_i x_i
\end{aligned}$$

[1]Note that for a set of cardinality n, there are $2^n = \sum_{i=0}^{n} \binom{n}{i}$ possible subsets, including the empty set.

As would be expected, $\varphi_i = w_i x_i$, and hence the Shapley decomposition recovers the correct effect.

The value of the Shapley decomposition emerges when the function is non-linear, as otherwise the decomposition is trivial. Consider a simple function that includes an interaction effect:

$$v(\{A,B\}) = a + b + ab,$$

where factor $A = 1$ maps to the effect of a, and factor $B = 2$ maps to the effect of b. Again, define $v(\varnothing) = 0$, $v(A) = a$ and $v(B) = b$. The assumption is here that absent factors are set to values of zero. The contributions are then:

$$\begin{aligned}
\varphi_A &= \tfrac{1}{2}\left[v(\{A,B\}) - v(A)\right] + \tfrac{1}{2}\left[v(A) - v(\varnothing)\right] \\
&= \tfrac{1}{2}(a + b + ab - b) + \tfrac{1}{2}a \\
&= a + \tfrac{1}{2}ab \\
\varphi_B &= b + \tfrac{1}{2}ab
\end{aligned}$$

The Shapley decomposition thus distributes the interaction term equally among the constitutive factors. In the absence of other information, this solution makes intuitive sense.

4.4 Properties of the Shapley value

As a *general* decomposition rule, the Shapley has a number of desirable properties. These will be listed briefly here. For a more rigorous treatment, see Joseph (2020), Shapley (1953), Strumbelj and Kononenko (2010), and Young (1985).

1. **Path Independence**: Let i and j be two factors, and v and v' be identical value functions, except that in v', i has been labelled j and j has been labelled i. Then $\varphi_i(N,v) = \varphi_j(N,v')$. This shows that the contributions are invariant under relabeling of the factors, which is equivalent to path independence.

2. **Efficiency**: Let v be a value function defined for the factors $N = \{1,2,\ldots,m\}$. Then

$$\sum_{i=1}^{m} \varphi_i(N, v) = I.$$

This is the key property of the Shapley decomposition. A proof is found in the Appendix.

3. **Dummy Factor**: If $v(\gamma \cup \{i\}) = v(\gamma)$ for all $\gamma \subseteq N \setminus \{i\}$, then $\varphi_i = 0$. Hence, a factor that has marginal contributions that are always zero also has a total contribution of zero. This directly follows from (4.2).

4. **Symmetry/Substitution**: If $v(\gamma \cup \{i\}) = v(\gamma \cup \{j\})$ for all $\gamma \subseteq N \setminus \{i, j\}$, then $\varphi_i = \varphi_j$. Hence, two factors that have identical marginal contributions also have identical total contributions.

5. **Linearity**: As shown above, when $v(\cdot)$ is a linearly additive function of the form, $v(\gamma) = \sum_{i \in \gamma} w_i x_i$ for all $\gamma \subseteq N$, then $\varphi_i = w_i x_i$.

6. **Monotonicity**: Let v and v' be two value functions defined on N. If $v(\gamma \cup \{i\}) - v(i) \geq v'(\gamma \cup \{i\}) - v'(i)$ for all $\gamma \subseteq N \setminus \{i\}$, then $\varphi_i(N, v) \geq \varphi_i(N, v')$. Hence, a factor's total contribution does not decrease if the factor's marginal contributions increase or stay the same, irregardless of the other factors.

As Young (1985) shows, the Shapley value is the unique symmetric allocation procedure that is monotonic. This makes a strong case for the Shapley value as a general decomposition rule.

4.5 Algorithm

A major downside of the Shapley decomposition is its computational complexity. Usually, with a reasonably large number of factors, computing $v(\cdot)$ on a given set of factors takes much more time than computing the marginal differences. Given that all subsets of N have to be computed (including $v(N)$ and $v(\varnothing)$), there are 2^m computations necessary. Depending on the complexity of v, then, computing the Shapley decomposition for a large number of factors is often not feasible. I therefore implement a simple algorithm to approximate the solution of the Shapley decomposition by sampling randomly from the $m!$ elimination sequences.

Key to this algorithm is the fact that the weighting factor in (4.2), $|\gamma_j|!(m-1-|\gamma_j|)!$, ensures that subsets of different sizes are given equal weight. If there are m factors in total, there are 2^{m-1} subsets that exclude i. Of these subsets, there is $\binom{m-1}{0} = 1$ subset of size 0, there are $\binom{m-1}{1} = m-1$ subsets of size 1, there are $\binom{m-1}{2}$ subsets of size 2, and so on. It follows then, that for a subset of size w, the total weight given to the subsets of this size is

$$\binom{m-1}{w} w!(m-1-w)! = (m-1)!.$$

As this number does not depend on w, all subsets of different sizes are given equal weight. This fact suggests the use of a two stage algorithm: First, randomly choose w from $(0, 1, \ldots, m-1)$ with equal probability. Second, randomly choose a subset of size w with equal probability.

The algorithm requires two parameters: t is the minimum number of iterations, and s is the desired standard error. The algorithm to approximate φ_i for a factor i is as follows:

1. Repeat the following steps for $j = 1, 2, \ldots$
   (a) Sample an integer between 0 and $m-1$, call it w.
   (b) Sample w elements from $N \setminus \{i\}$ without replacement, call the resulting set R.
   (c) Calculate $\hat{\varphi}_i^j = v(R \cup \{i\}) - v(R)$.
   (d) If $j > t$, calculate $\hat{s}$, the standard error of the $\hat{\varphi}_i^j$'s, and stop if $s < \hat{s}$.
2. Let $\hat{\varphi}_i \equiv \frac{1}{M} \sum_{j=1}^{M} \hat{\varphi}_i^j$ where M is the number of $\hat{\varphi}_i^j$ sampled.

The number of contributions that are sampled is determined by t, the *minimum* number of contributions that are sampled for each factor, and s, the minimum standard error that is desired. After some experimentation, I set $t = 25$ and $s = 0.01$ for applications in the following two chapters.

4.6 Defining v and N

In any practical application of the Shapley decomposition, the key task of the researcher is to define v and N. In game theory, v and N are usually given, and the goal is to analyze the game. A classic example is the "glove game" (or, more generally, a market with a different number of buyers and sellers) (Aumann 1994). Assume three "players" (in the game-theoretic sense), where players 1 and 2 hold a left-hand glove, and player 3 holds a right-hand glove. The goal is to form coalitions where both a left-hand and a right-hand glove are present. Then, $N = \{1, 2, 3\}$ and

$$v(\gamma) = \begin{cases} 1 & \text{if } \{3\} \in \gamma \text{ and } |\gamma| \geq 2, \\ 0 & \text{otherwise,} \end{cases}$$

where $\gamma \subseteq N$. The Shapley values for this game are $\varphi_1 = \varphi_2 = 1/6$ and $\varphi_3 = 2/3$. Player 3's higher value reflects the fact that they hold a scarce resource.

For decomposition in the social sciences, however, giving meaning to the "players," i.e. the factors of the decomposition, as well as the setup of the function v, are itself part of the research design. To illustrate, consider a classic setup from Oaxaca-Blinder-Kitagawa regression decomposition, where two separate linear regression functions have been estimated for two different subgroups. Let $\hat{\boldsymbol{\beta}}_1$ and $\mathbf{X}_1$ denote the vector of regression coefficients and the data matrix for the first subgroup, respectively, and $\hat{\boldsymbol{\beta}}_2$ and $\mathbf{X}_2$ the same for the second subgroup. Also define the vectors $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ containing the sample means of $\mathbf{X}_1$ and $\mathbf{X}_2$. The goal of regression decomposition is to attribute the difference in the predicted means, $\bar{\hat{y}}_2 - \bar{\hat{y}}_1 = \bar{\mathbf{x}}_2^T \hat{\boldsymbol{\beta}}_2 - \bar{\mathbf{x}}_1^T \hat{\boldsymbol{\beta}}_1$ into two factors: One factor accounts for group differences between the predictor means $\bar{\mathbf{x}}$ (called the "endowment" effect), and one factor accounts for differences in the coefficients $\hat{\boldsymbol{\beta}}$.

If we let player $E = 1$ represent the endowment effect, and player $C = 2$ represent the coefficients effects, then

$$
\begin{aligned}
v(N) = v(\{E, C\}) &= \bar{\mathbf{x}}_2^T \hat{\boldsymbol{\beta}}_2 = \bar{\hat{y}}_2, \\
v(E) &= \bar{\mathbf{x}}_2^T \hat{\boldsymbol{\beta}}_1, \\
v(C) &= \bar{\mathbf{x}}_1^T \hat{\boldsymbol{\beta}}_2, \\
v(\varnothing) &= \bar{\mathbf{x}}_1^T \hat{\boldsymbol{\beta}}_1 = \bar{\hat{y}}_1,
\end{aligned}
$$

such that $I = v(N) - v(\varnothing) = \bar{\hat{y}}_2 - \bar{\hat{y}}_1$, the outcome of interest. As there are only two factors, the Shapley contributions again have a simple form:

$$
\begin{aligned}
\varphi_E &= \frac{(\hat{\boldsymbol{\beta}}_1 + \hat{\boldsymbol{\beta}}_2)^T}{2} (\bar{\mathbf{x}}_2 - \bar{\mathbf{x}}_1), \\
\varphi_C &= \frac{(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)^T}{2} \left(\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1\right).
\end{aligned}
\tag{4.3}
$$

This solution is identical to the results of Kitagawa (1955) and Reimers (1983). It has the advantage that the choice of reference group is not relevant, due to the symmetry property of the Shapley value. Note also that in this formulation, the grand mean of $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ is used in (4.3), while other formulations use a weighted mean of the two coefficient vectors (see Jann 2008). From the perspective of the Shapley value, such weighted means formulations are not permitted.

The Shapley decomposition reveals the key assumption of the Oaxaca-Blinder-Kitagawa regression decomposition, which is that the quantities $v(E)$ and $v(C)$ are defined sensibly. Except for $v(N)$ and $v(\varnothing)$, which are given as data, all other values of $v(\cdot)$ are *counterfactuals*. The counterfactual $v(E)$ determines what the mean predicted value would be if we combined the means of the second subgroup with the coefficients of the first subgroup. Whether this is a reasonable counterfactual, and especially whether it can be given a causal attribution, depends on factors external to the decomposition rule. The transparency with which the assumptions of the decomposition can be assessed directly from the form of the counterfactuals $v(\cdot)$ is an advantage of the Shapley decomposition over more specialized decomposition rules.

Starting from this framework, it is also straightforward to extend the decomposition to other factors that are of interest. For instance, instead of focusing on the aggregate effect C, one could decompose this effect further into the effect of individual predictors, as routinely done in regression decomposition. In the Shapley framework, this amounts to modifications of the set of factors N, which is now $N' = \{E, C_1, \ldots, C_p\}$ (assuming p predictors), and the value function v, which now has to provide a value for any subset of N'. Again, the task of the researcher is to define appropriate counterfactuals of the form $v(\gamma)$ where $\gamma \subseteq N'$.

4.7 Decomposition of segregation change

In the following two chapters, I apply the Shapley decomposition to study segregation change. The general setup is as follows. Let $\mathbf{T}_1$ be the relevant contingency table at time point 1, and $\mathbf{T}_2$ be the contingency table at time point 2. The remaining chapters make use only of the H index, so write $H(\mathbf{T}_1)$ for the segregation index at time point 1, and $H(\mathbf{T}_2)$ for the segregation index at time point 2. The outcome of interest is then

$$I = H(\mathbf{T}_2) - H(\mathbf{T}_1),$$

such that $H(\mathbf{T}_2) = v(N)$ and $H(\mathbf{T}_2) = v(\varnothing)$.

One simple decomposition is to attribute changes in segregation to each of the racial groups. For the case of two racial groups, let the set of factors be $N = \{A, B\}$, where $A = 1$ stands for the impact of population A on segregation change, and $B = 2$ stands for the impact of the population B on segregation change. When both factors are included, the H value at time point 2 is obtained, and if no factors are included, the H at time point 1 is obtained. The remaining counterfactuals, $v(A)$ and $v(B)$, are defined as follows: For $v(A)$, I calculate the H index for a matrix where the racial group counts for population A come from $\mathbf{T}_2$, but the racial group counts for population B come from $\mathbf{T}_1$. Thus, the index obtained through $v(A)$ reflects a hypothetical situation where population B remains in place, while the population A is distributed as in $\mathbf{T}_2$. The counterfactual $v(B)$ is defined equivalently.

As a simple example, consider the following two contingency tables, where the first column contains counts for population A, and the second column contains counts for population B. This very small metropolitan area contains just three spatial units (e.g., neighborhoods):

$$\mathbf{T}_1 = \begin{bmatrix} 10 & 40 \\ 10 & 40 \\ 80 & 20 \end{bmatrix} \qquad \mathbf{T}_2 = \begin{bmatrix} 20 & 35 \\ 20 & 35 \\ 60 & 30 \end{bmatrix}$$

The H indices for these tables are $H(\mathbf{T}_1) = 28$ and $H(\mathbf{T}_2) = 6$, for a decline of $I = -22$. Just from visual inspection of the matrices, it seems that population A redistributed more thoroughly. The first two neighborhoods gained 10 members of population A and lost five members of population B, while the third neighborhood lost 20 members of population A, and gained 10 members of population B. Both groups are now more evenly distributed across units, but how much impact had each racial group on the total reduction in segregation? Let

$$v(A) = H(\begin{bmatrix} 20 & 40 \\ 20 & 40 \\ 60 & 20 \end{bmatrix}) = 12,$$

where the first column is taken from $\mathbf{T}_2$, and the second column from $\mathbf{T}_1$. This reflects the counterfactual situation that only population A has redistributed. Also,

$$v(B) = H(\begin{bmatrix} 10 & 35 \\ 10 & 35 \\ 80 & 30 \end{bmatrix}) = 19,$$

where the second column is taken from $\mathbf{T}_2$, and the first column from $\mathbf{T}_1$.

Then compute

$$\begin{aligned} \varphi_A &= 1/2\,[v(\{A,B\}) - v(B)] + 1/2\,[v(A) - v(\varnothing)] \\ &= 1/2\,[6 - 19] + 1/2\,[12 - 28] = -14.5 \end{aligned}$$

and

$$\begin{aligned} \varphi_B &= 1/2\,[v(\{B,A\}) - v(A)] + 1/2\,[v(B) - v(\varnothing)] \\ &= 1/2\,[6 - 12] + 1/2\,[19 - 28] = -7.5 \end{aligned}$$

For this simple example, we would therefore conclude that about 65% of the decline can be attributed to changes in the distribution of population A.

At this point, the decomposition can be extended in various ways. In the following two chapters, I will focus on decompositions that distinguish between neighborhoods based on their racial composition at time point 1, and the growth trajectory of each population. In the example, population A grew in majority-B neighborhoods, and declined in majority-A neighborhoods. These two processes, of course, should both lead to declining segregation. However, again, the question is how much these processes contribute to the total decline. When classifying each neighborhood as either majority-A or majority-B, there are eight decomposition factors:

1. growth of population A in majority-A neighborhoods,
2. decline of population A in majority-A neighborhoods,
3. growth of population B in majority-A neighborhoods,
4. decline of population B in majority-A neighborhoods,
5. growth of population B in majority-B neighborhoods,
6. decline of population B in majority-B neighborhoods,
7. growth of population A in majority-B neighborhoods,
8. decline of population A in majority-B neighborhoods.

Theoretically, we would expect that factors 2, 3, 6, and 7 are associated with declines in segregation, while the remaining factors are likely associated with increases in segregation. I will make use of a similar decomposition in the following two chapters. Importantly, this decomposition will allow me to answer such questions as to whether segregation declines in the U.S. have been brought about mainly by the White population moving to minority neighborhoods, or vice versa.

It is possible to increase the number of factors further and distinguish neighborhoods, for instance, by income levels, population density, or other locational factors. The choices for the factors N and value function v are endless, and should be informed by theoretical considerations.

4.8 Conclusion

This chapter has formally introduced the Shapley decomposition, and presented some of its desirable properties. The Shapley decomposition is a rule that assign contributions to each factor for any value function v defined on the set of factors N. In the social sciences, defining v and N are the key tasks of the researcher, which mainly involves defining the appropriate counterfactuals $v(\gamma)$ for all $\gamma \subseteq N$. Lastly, it has been shown by example how the Shapley decomposition can be used to decompose changes in segregation.

4.9 Appendix: Proof of Efficiency

Let $\Gamma_w(\cdot)$ be the subsets of size w of a set. For instance, if $N = \{1,2,3\}$, then $\Gamma_3(N) = \{\{1,2,3\}\}$, $\Gamma_2(N) = \{\{1,2\},\{2,3\},\{1,3\}\}$, $\Gamma_1(N) = \{\{1\},\{2\},\{3\}\}$, and $\Gamma_0(N) = \varnothing$.

Then derive a telescoping series from eq. (4.2):

$$\begin{aligned}
\sum_{i=1}^{m} \varphi_i &= \sum_{i=1}^{m} \frac{1}{m!} \sum_{\gamma_j \subseteq N\setminus\{i\}} |\gamma_j|!(m-1-|\gamma_j|)!\big(v(\gamma_j \cup \{i\}) - v(\gamma_j)\big) \\
&= \sum_{i=1}^{m} \frac{1}{m!} \sum_{w=0}^{m-1} \sum_{\gamma_j \in \Gamma_w(N\setminus\{i\})} w!(m-1-w)!\big(v(\gamma_j \cup \{i\}) - v(\gamma_j)\big) \\
&= \sum_{w=0}^{m-1} \frac{w!(m-1-w)!}{m!} \sum_{i=1}^{m} \left(\sum_{\gamma_j \in \Gamma_w(N\setminus\{i\})} v(\gamma_j \cup \{i\}) - \sum_{\gamma_j \in \Gamma_w(N\setminus\{i\})} v(\gamma_j) \right) \\
&= \sum_{w=0}^{m-1} \frac{w!(m-1-w)!}{m!} \left(\sum_{\gamma_j \in \Gamma_{w+1}(N)} (w+1)\, v(\gamma_j) - \sum_{\gamma_j \in \Gamma_w(N)} (m-w)\, v(\gamma_j) \right) \\
&= \sum_{w=0}^{m-1} \left(\sum_{\gamma_j \in \Gamma_{w+1}(N)} \binom{m}{w+1}^{-1} v(\gamma_j) - \sum_{\gamma_j \in \Gamma_w(N)} \binom{m}{w}^{-1} v(\gamma_j) \right) \\
&= v(N) + \sum_{w=1}^{m-1} \left(\sum_{\gamma_j \in \Gamma_w(N)} \binom{m}{w}^{-1} v(\gamma_j) - \sum_{\gamma_j \in \Gamma_w(N)} \binom{m}{w}^{-1} v(\gamma_j) \right) - v(\varnothing) \\
&= v(N) - v(\varnothing) \\
&= I
\end{aligned}$$

This development relies on the fact that, for a given w,

$$\sum_{i=1}^{m} \sum_{\gamma_j \in \Gamma_w(N \setminus \{i\})} v(\gamma_j \cup \{i\}) = \sum_{\gamma_j \in \Gamma_{w+1}(N)} (w+1)\, v(\gamma_j).$$

To see why this is true, note that there are $m|\Gamma_w(N \setminus \{i\})| = m\binom{k}{w} = \frac{m!}{w!(k-w)!}$ terms on the left-hand-side, and that the right-hand side represents the same number of terms, namely $|\Gamma_{w+1}(N)(w+1)| = \binom{m}{w+1}(w+1) = \frac{m!}{w!(k-w)!}$.

The second relevant equality is similar:

$$\sum_{i=1}^{m} \sum_{\gamma_j \in \Gamma_w(N \setminus \{i\})} v(\gamma_j) = \sum_{\gamma_j \in \Gamma_w(N)} (m-w)\, v(\gamma_j)$$

Again, there are $m|\Gamma_w(N \setminus \{i\})| = \frac{m!}{w!(k-w)!}$ terms on the left-hand side, and $|\Gamma_w(N)|(m-w) = \binom{m}{w}(m-w) = \frac{m!}{w!(k-w)!}$ terms on the right-hand side.

Chapter 5: Explaining Long-Term Changes in Residential Racial Segregation: A Demographic Decomposition Approach

5.1 Introduction

The changing racial group distribution in the U.S., sometimes referred to as a “diversity explosion” (Frey 2018), has attracted the interest of many scholars in sociology, demography, and economics. The diversity explosion is mostly a product of immigration, especially of “new minorities,” such as Hispanics and Asians, alongside other immigrant groups. However, other demographic factors play a role as well. The White population is rapidly aging, and its growth rate has stalled. In addition, metropolitan areas in the South and West of the U.S. experience population growth, both through immigration and by attracting migration from Northern metropolitan areas. These large-scale demographic trends have prompted a large literature on the impact of these changes on political polarization, labor relations, and marital homogamy, among other topics.

Of special interest is how these trends relate to declines in residential segregation. Especially for the Black population, segregation has steadily declined since the 1970s. For Hispanics and Asians, segregation is usually reported as either stable or slightly increasing. Rarely, however, have studies attempted to draw a direct link from changes in residential diversity to changes in racial segregation. This paper fills this gap by studying changes in U.S. residential segregation between 1990 and 2010 for the 224 largest metropolitan areas. The core contribution of this paper lies in the application of a flexible decomposition method that allows us to understand which racial groups and spatial units in a metropolitan area contributed to segregation decline or increase.

The current study takes an explicit population-based approach to explaining segregation change. Many empirical studies of segregation change seek to explain the aggregate segregation index of a metropolitan area through aggregate measures, without taking into account the population movements that are required to engender change. To take a simple example, if segregation between Blacks and Whites declined, this can reflect two distinct processes: Either Blacks move into majority White areas, or Whites move into majority Black areas. We know that both of these processes are occurring—the former under the label of "Black suburbanization," the latter as "gentrification"—, however, we do not know the relative importance of each of these processes in bringing about changes in segregation.

The goal of this study is therefore (1) to identify the relevant population processes that shape neighborhood change, and (2) to quantify how much these processes contribute to changes in segregation. In this "segregation-as-process" perspective, I shift the emphasis from the aggregate segregation measure to the mechanical demographic processes that shape neighborhood mobility, and therefore segregation. Population movements (and their absence) may be considered as the "proximate causes" for segregation change, compared to the remote causes that affect population movements (or their absence).

This study builds on earlier results in segregation measurement that have shown the advantages of decomposable segregation indices. Throughout this study, I use the entropy-based segregation index H (ranging from 0 to 100). In line with earlier research (e.g., Farrell 2008; Lichter et al. 2015), segregation is decomposed into macro and micro components. Macro segregation refers to place-based segregation. Places, as local governmental units, include the principal city or cities of a metropolitan area, as well as incorporated places in the suburbs and exurbs. I therefore refer to block-level segregation in a metropolitan area as *total segregation*, to segregation between places as *macro segregation*, and to block-level segregation within places as *micro segregation*. For a given metropolitan area, the three quantities are related through the simple identity *macro* + *micro* = *total*.

The remainder of the paper is organized as follows. First, findings about the empirical trends in macro and micro segregation are reviewed. Special attention is paid to the methods that seek to explain *changes* in segregation. The common method of regressing segregation indices on metropolitan area attributes is found to be lacking, as it may not succeed in identifying the significant drivers of segregation change. Instead, a decomposition method is proposed that directly identifies the fundamental demographic factors that produce segregation change, such as population growth and decline of the different racial groups in various parts of the metropolitan area. The decomposition results show, for the first time, that stable macro segregation is a product of countervailing forces: While minority groups move towards integration, White population movements increase segregation. These dynamics are fundamentally shaped by the relationship between cities and suburbs; minority groups have suburbanized in great numbers, but White population declines in principal cities and integrated suburbs counteract such trends. While macro segregation is effectively unchanged over the period under study, micro segregation has declined. I find that these declines are similarly driven by minority groups, however, counteracting forces are less prevalent.

The upshot of these findings is that most metropolitan areas are shaped by simultaneous and ongoing *de*segregation and *re*segregation. For macro segregation, Black integration and White resegregation offset each other, leading to unchanged macro segregation in total. This shows that the absence of change in the aggregate segregation index does not reflect the absence of neighborhood change or residential mobility. For micro segregation, declines in segregation produced by Black integration are not offset by White resegregation, leading to declines in total metropolitan area segregation.

5.2 Background

5.2.1 Hierarchies of segregation

Studies of residential segregation routinely distinguish segregation at different geographic scales. A standard approach is to focus on block- or tract-level segregation within a metropolitan area, and then distinguish sources of segregation that arise at different levels of the spatial hierarchy. For instance, studies may quantify how much of metropolitan area segregation is due to segregation between places, compared to segregation within places. Consistent with the terminology in the literature, I will refer to segregation at smaller scales, such as tracts or blocks, as "micro" segregation, and segregation at higher scales, such as places, communities, or municipalities, as "macro" segregation (Farrell 2008; Lichter et al. 2015).

Theoretically, it is possible that all segregation is either entirely between places *or* entirely within places. In the first case, places are internally racially homogeneous, but differ in their racial composition from each other (for instance, suburbs could be all-White, while central cities could be all-Black). In the second case, places are internally segregated, but all contain the same proportions of each racial group. Empirically, both between-place and within-place segregation will be present, but their relative weights are indicative of different forms of segregation (Farrell 2008; Lichter et al. 2015).

If within-place segregation is dominant, segregation is a localized phenomenon that mostly operates between neighborhoods. Neighborhoods characteristics, such as the types of homes available (single family or apartments, public housing), provide opportunities for some racial groups, and impose restrictions for others. This situation is compatible with the place stratification model of neighborhood segregation, where neighborhoods are embedded in a hierarchy of desirable locations. While many members of minority groups have the resources to move to a better neighborhood, institutional and individual forces (such as the real estate market, housing discrimination, family networks) preserve the racial composition of neighborhoods. An important driver of maintaining neighborhood boundaries is housing discrimination, which was likely the main reason for high levels of segregation observed before 1970 (Charles 2003). Glaeser and Vigdor (2012) argue that the end of legal housing discrimination (e.g., Fair Housing Act of 1968) has contributed to declines in segregation, but many studies find that housing discrimination and racial prejudice persists (for a review, see Charles 2003; Korver-Glenn 2018).

On the other hand, high place-based segregation will be dominant when places reflect political entities that enact exclusionary policies. Places, in this sense, are governmental units with the political and economic power to enact legislation and policy to consciously or unconsciously exclude certain groups. An extreme example of a place-based exclusionary policy are "sundown towns," where members of a minority group (usually African Americans) were legally required to leave the town before sunset (Loewen 2018). Usually, the exclusionary effects are more subtle, reflected by policies such as low-density zoning, restrictions on building activity, or racially-biased annexation (Pendall 2000; Rothwell and Massey 2010). Clearly, the place stratification model also applies to macro segregation. In this case, the hierarchy of neighborhoods and places is further reinforced through legal actions.

5.2.2 Empirical trends in micro-macro segregation

The literature is unanimous in its conclusion that average residential racial segregation has decreased since the 1970s. Logan et al. (2004) find that Black-White segregation has decreased by 12% between 1980 and 2000, while Hispanic-White and Asian-White segregation has slightly increased—although starting from much lower absolute levels compared to Black-White segregation. Other studies using different measures find similar trends (Charles 2003; Farley and Frey 1994; Iceland 2004; Lichter et al. 2015; Reardon et al. 2009). These patterns differ by metropolitan area: declines in segregation have been concentrated in areas where segregation was already low (such as the South and the West), and many large metropolitan areas in the Northeast and Midwest continue to be characterized by "hyper-segregation" (Charles 2003; Iceland 2004; Massey and Denton 1989). These findings are robust to the use of different segregation indices, different definitions of metropolitan areas and sample inclusion criteria, and different scales at which segregation was measured (Census tract, block group, or block).

With regards to the trends in macro and micro segregation, findings are more divided. Massey et al. (2009) study Black-White segregation during the entire 20th century. They show tract-level (micro) segregation increased until about 1970, and then started to decline.

Segregation between cities increased together with tract-level segregation until about 1980, but has since remained flat. The study of Fischer et al. (2004), for the period 1960-2000, is among the earliest papers using a macro-micro decomposition. Their study shows that declining residential segregation is mostly driven by micro segregation—especially for Blacks. Fischer (2008) further finds that declines in segregation between Blacks and Whites are mostly due to declining central city segregation, although declines in segregation were also observed in the suburbs.

Reardon et al. (2009) (see also Lee et al. 2008) deviate from the usual approach of using Census geographies (such as places or tracts) directly, and instead compute spatially-weighted segregation indices. Despite the different methodology, the results are broadly in line with similar studies: Black-White segregation declined at the micro level, but stayed stable at the macro level. Hispanic-White and Asian-White micro and macro segregation both increased, but from a much lower level than Black-White segregation.

The studies by Farrell (2008) and especially Lichter et al. (2015) argue most prominently that macro segregation is gaining new importance. For instance, Farrell (2008, p. 467) writes that "urban and suburban municipalities are replacing neighborhoods as the central organizing units of metropolitan segregation." Lichter et al. (2015) seem to provide the most compelling evidence, and they claim that there exists "a new macro-segregation, where the locus of racial differentiation resides increasingly in socio-spatial processes at the community or place level" (p. 843).

I argue that the shift towards macro segregation, as described in these studies, has been exaggerated. The results by Farrell (2008) support the finding that macro segregation has increased, but only slightly—especially when compared to the declines in micro segregation, which are an order of magnitude larger than the increases in macro segregation.[1] Below, I replicate the study by Lichter et al. (2015) and show that their own results are in fact in line with stable (Black-White), or *slightly* increasing (Hispanic-White, Asian-White) macro segregation.

[1] For instance, in Farrell (2008, Table 3), the total decline in micro segregation (within cities and within suburbs) is -0.0332, while the increase in macro segregation is 0.0026. (Here, the H is measured on a scale of 0 to 1.)

The finding that stands out with regard to segregation change is again a comparatively large decline in micro segregation.

5.2.3 Explaining changes in segregation

The declines in total segregation and micro segregation, as well as the (alleged) increases in macro segregation, call for explanations. The widespread decrease in micro segregation runs counter to the place stratification model. The model can still be useful to explain ongoing high levels of residential segregation and the, arguably, slow pace of progress, but it cannot explain the large declines in micro segregation. An alternative model is the spatial assimilation model, which posits that minority groups become residentially integrated as they are integrated along other dimensions, such as income or education. Members of minority groups therefore "move up" spatially when they move up in the distribution of other status characteristics.

The approach that is often taken to test theories such as the place stratification or spatial assimilation model is to estimate regression models at the level of the metropolitan area, where segregation scores or changes in segregation are regressed on characteristics of metropolitan areas. Commonly included predictors are population shares of the different racial groups, the size of the metropolitan area, its specialization (e.g., in manufacturing, or retirement or military communities), average household income for different racial groups, income differences between the racial groups, and the percentage of foreign-born residents.

This type of analysis has been pioneered by Farley and Frey (1994). Their models show that younger and smaller metropolitan areas in the South and West, with a lower proportion of black residents, generally have the lowest levels of segregation. Many later studies continue in this spirit. For instance, Logan et al. (2004), Lee et al. (2008), and Reardon et al. (2009) confirm Farley and Frey's findings for later periods. Iceland (2004), using a metropolitan-area fixed-effects model, finds that segregation increases with rising racial diversity.

What is common to the regression-model approach is that the explanatory power can only come from changes or characteristics at the level of metropolitan area. These explanations are therefore disconnected from the more localized explanations that are usually cited to explain segregation in the U.S. For instance, differences in socioeconomic status between racial groups are large in the U.S., and several studies find that metropolitan areas in which the income differences between racial groups are smaller tend to have lower segregation as well, consistent with spatial assimilation theory (Logan et al. 2004; Reardon et al. 2009). However, these differences cannot fully explain patterns of residential segregation, and this is especially true for Blacks, who remain more segregated than their economic status would predict (Galster and Sharkey 2017). Logan (2016) argues that the high levels of segregation still observed are the product of the perpetuation of settlement patterns. For instance, public housing projects in many urban areas continue to house overwhelmingly Black residents. Whites, on the other hand, with more wealth and more financial freedom, have accumulated housing stock in the suburbs that is passed on to the next generation. This suggests ongoing patterns of economic disadvantage that hinder Black families' exit of public housing, and/or the persistence of housing discrimination. While such arguments are intriguing, they are hard to test at the level of the metropolitan area. Furthermore, the predictors of these regressions are usually designed to reveal the factors that are associated with *increases* in segregation. The results are therefore often not helpful in understanding the declines in segregation that are empirically observed.

Two other disadvantages of the regression-model approach need to be mentioned. First, by regressing an aggregate segregation index on other aggregate measures, regression models at the level of the metropolitan area are possibly subject to the ecological fallacy. For instance, assume that wealthy individuals are more likely to segregate. Then it can still be the case that, of course, wealthy metro areas are less segregated. The relationship depends on the distribution of wealth, and if wealthy metro areas have higher income inequality, they will also be less segregated.

A second problem is that the regression model may not be successful in explaining the processes that we would like to understand. Assume for instance, that micro segregation declined everywhere at a similar rate, and that the decline is mostly explained by Black suburbanization.

As I show below, this scenario is not an entirely uncharacteristic representation of the actual situation. The problem is that there is not much variability in the "treatment" (Black suburbanization), such that the regression model will not be able to identify it as an important factor. As Bhrolcháin and Dyson (2007) put it, in such a situation the "true explanandum is the uniformity of decline rather than its variability" (p. 5).

5.3 Goal of the current study

The current study takes an explicit population-based approach to explaining segregation change. Many empirical studies of segregation change attempt to "explain" the aggregate segregation index of a metropolitan area through aggregate measures, without taking into account the population movements that are required to bring about change. To take a simple example, if micro segregation between Blacks and Whites declined, this can reflect two distinct processes: Either Blacks move into majority White areas, or Whites move into majority Black areas. We know that both of these processes are occurring, however, we do not know the relative importance of each of these processes in bringing about segregation change.

A similar argument applies to segregation decomposition in the cross-section, as discussed above. Recent studies have often decomposed total metropolitan area segregation into the sources of segregation at different scales, for instance in components that reflect how much segregation is due to the suburbs compared to the central city. These decompositions are essential in localizing the sources of segregation, and, when compared at different points in time, they also hint at possible explanations. However, these decompositions are still disconnected from the underlying population movements that produced this change.

The goal of this study is therefore (1) to identify the relevant population processes that shape neighborhood change, and (2) to quantify how much these processes contribute to changes in segregation. In this "segregation-as-process" perspective, I shift the emphasis from the aggregate segregation measure to the mechanical demographic processes that shape neighborhood mobility, and therefore segregation. Population movements (and their absence) may be considered as the "proximate causes" for segregation change, compared to the remote causes that affect population movements (or their absence).

I give here some examples of problems that seem fundamental, but are hard to engage with in the regression framework:

- Common predictors in ecological regression models are racial group shares, such as the share of the Black population. A regression coefficient for a proportion, however, is hard to interpret: The share of the Black population could increase because Blacks are moving in, but also because other racial groups move out. By studying the underlying population movements, these two dynamics can be distinguished, and a fundamental question of segregation change be answered: Are declines in segregation driven by the mobility of the White or the Black population?

- The proposed perspective also allows us to more directly engage with large-scale population trends that have been reported in the literature, such as the "diversity explosion" (Frey 2018). Importantly, changes in diversity and racial composition do not have to lead to changes in segregation (Fowler et al. 2016). Furthermore some processes might have different impacts in different metropolitan areas: immigration could lead to both an increase or a decrease in segregation, depending on local circumstances (Logan et al. 2004).

- As a one-number summary, a segregation index can also obscure countervailing trends. For instance, zero change in the segregation index can be accounted for either by no population movement (rather unlikely), or by offsetting movements. For instance, white flight and Black suburbanization could happen simultaneously in different parts of the metropolitan area, cancelling each other out. A similar offsetting argument can be made for groups that are composed both of newly arrived immigrants and native-born residents. While immigrants cluster in certain neighborhoods, long-term residents may be able to "move up" in the spatial hierarchy. This is the case for Asians and Hispanics (Hall

> 2013). Importantly, then, while the net segregation change that these dynamics produce may be zero, this does not indicate the absence of change.

The number of studies that have directly connected population mobility to segregation change is small. This is likely due both to data constraints and the absence of a methodology to decompose changes in segregation. Two exceptions to this are the studies by Winkler and Johnson (2016) and Bader and Warkentien (2016). The latter distinguish between three kinds of segregation trajectories at the neighborhood level: durable segregation, racial change, and durable integration. They argue convincingly that segregation should not be studied in the cross-section, because this gives only a snapshot of what is ultimately a segregation *process*. The snapshot might capture an intermediate, more integrated picture, although the trajectory points towards resegregation. One example of this is "white flight": A neighborhood were Blacks have started to move in and Whites have started to move out might look integrated when observed in the middle of this process. However, Bader and Warkentien (2016) do not connect their findings to aggregate segregation measures in the metropolitan areas they study.

The approach that Winkler and Johnson (2016) pursue is similar to the research design of the current paper. The authors show that between 1990 and 2010, between-county migration served to decrease segregation, while natural population change would have increased segregation. They also report age differences: The mobility of young people (below 40) reduced segregation, while the mobility of people over age 60 increased segregation. With this approach, Winkler and Johnson (2016) go far beyond usual studies on segregation, but the paper is not without limitations: Significantly, due to data constraints, the smallest unit of analysis is only the county, which excludes a large—and arguably the most significant—amount of residential segregation. (Note that one county in a dense metropolitan area may include dozens of places, and thousands of blocks.) Secondly, while the authors compute counterfactual segregation indices, not all their results are full decompositions in the sense that the sum of all factors reflects the total change in segregation.

5.4 Data and Methods

In studies of segregation, two research designs are commonly used: cross-sectional and harmonized studies. In a cross-sectional study, data is collected (usually for different Census years), and the geographical boundaries that are valid in each year are used. This is relevant for both blocks and places: Blocks are redrawn for each Census (although many are stable), and places may have expanded or contracted. This cross-sectional approach was used, for instance, by Lichter et al. (2015). In this paper, however, I am interested in decomposing changes in segregation by quantifying the impact of population growth and decline in certain, fixed areas. This design requires stable geographies, and I am therefore using a harmonized design, where 2010 blocks definitions are applied to 1990 boundaries. The crosswalks are provided by IPUMS NHGIS (Manson et al. 2019).[2] The crosswalk will introduce uncertainty into the estimates, and I compare results from the cross-sectional and harmonized files to show that this uncertainty is small.

The full dataset is constructed as follows, closely replicating the design used by Lichter et al. (2015). I obtain block-level racial group counts from the Census datasets for 1990, 2000, and 2010 through IPUMS NHGIS (Manson et al. 2019). I apply the 2013 metropolitan area definitions from the Office of Management and Budget[3] to these files, and remove metropolitan areas that contain less than 1,000 people of any of the four major racial groups (White, Black, Hispanic, Asian). These four racial groups are coded as mutually exclusive (i.e., "White" refers to "non-Hispanic White"). This leaves 224 metropolitan areas. Blocks nest perfectly in places, and each block can be attributed uniquely toward a place or a non-place area. In line with earlier research, I refer to non-place areas, including Census-designated places, as "fringe areas."[4] To construct the harmonized file, I crosswalk 1990 and 2000 block definitions toward 2010 block

[2] For details on the procedure, see `https://www.nhgis.org/user-resources/geographic-crosswalks`.

[3] Available from `https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html`.

[4] I am not aware that other studies that distinguish between place and non-place areas include Census-designated places in the non-place category. I argue that this approach is more sensible, as the interest in studying places is usually prompted by their legal status. As the name implies, Census-designated places do not coincide with any governmental or legal function, and I therefore choose to include them as "fringe areas."

Table 5.1: Average racial and spatial composition of metropolitan areas

	1990			2010		
	25 pct.	Mean	75 pct.	25 pct.	Mean	75 pct.
Racial Composition (in %)						
White	63	79	85	49	68	74
Black	6	11	18	7	11	17
Hispanic	2	8	15	6	14	27
Asian	1	2	4	2	4	7
Spatial Composition (in %)						
Principal city	30	42	51	26	39	46
Suburban places	17	22	39	18	23	40
Fringe	14	35	45	16	38	47

Note: The quantities shown are population-weighted percentiles.
Example: In 1990, 50% of the U.S. metropolitan population lived in metropolitan areas that were between 1% and 4% Asian.

definitions, and then apply place definitions as of 2010. Therefore, the number of blocks and places is stable in the harmonized data set. In the cross-sectional data set, the number of blocks and places will be different in each year, and, as the U.S. population has grown between 1990 and 2010, the number of blocks and places will have increased in most metropolitan areas.

An overview of the data set is shown in Table 5.2, which displays the racial and spatial composition of the 224 metro areas, using the harmonized samples. Shown are the population-weighted 25th, 50th, and 75th percentiles for 1990 and 2010.

5.4.1 Segregation measure

Throughout this study, I use the entropy-based segregation index H. To define H in general terms, let $\mathbf{T}$ be a matrix with U rows (spatial units) and G columns (racial groups). The rows U represent spatial units within a metropolitan area or a subarea, e.g. blocks or places. Let the entries of $\mathbf{T}$ be t_{ug}, the number of people of race g in spatial unit u, and let t be the total population of $\mathbf{T}$, i.e. $t = \sum_{u=1}^{U}\sum_{g=1}^{G} t_{ug}$. The joint probability of being in spatial unit u and racial group g is $p_{ug} = t_{ug}/t$. Also define $p_{u\cdot} = \sum_{g=1}^{G} t_{ug}/t$ and $p_{\cdot g} = \sum_{u=1}^{U} t_{ug}/t$ as the marginal probabilities of spatial units and racial groups, respectively. The index H is then defined as

$$H(\mathbf{T}) = \frac{100}{E(\mathbf{T})} \sum_u \sum_g p_{ug} \log \frac{p_{ug}}{p_{u\cdot} p_{\cdot g}},$$

where $E(\mathbf{T}) = -\sum_{g=1}^{G} p_{\cdot g} \log p_{\cdot g}$ is the entropy of the racial group marginal distribution of $\mathbf{T}$. In this formulation, the index ranges from 0 (absence of segregation) to 100 (complete segregation).

Using this general formulation, it is possible to calculate a number of H indices for a given metropolitan area. To quantify segregation in the entire metro area, define the matrix $\mathbf{B}$, which contains as rows all Census blocks that belong to the metro area. The result of $H(\mathbf{B})$ will then quantify block-level segregation in the entire metropolitan area, referred to here as *total segregation.*

To make use of the micro-macro decomposition, also define a matrix $\mathbf{P}$ that aggregates Census blocks to places. This is possible because each block uniquely belongs to a place or a non-place area. Note that matrices $\mathbf{B}$ and $\mathbf{P}$ describe the same population, but that $\mathbf{P}$ contains many fewer rows than $\mathbf{B}$. Further, define a matrix $\mathbf{B}_s$, which contains the subset of blocks that belong to place s. If we stack all matrices $\mathbf{B}_s$ for a given metropolitan area, we obtain matrix $\mathbf{B}$ again. The decomposition of $H(\mathbf{B})$ is then given by

$$\begin{aligned}
H(\mathbf{B}) &= H_{\text{macro}}(\mathbf{B}) + H_{\text{micro}}(\mathbf{B}) \quad \text{where} \\
H_{\text{macro}}(\mathbf{B}) &= H(\mathbf{P}) \\
H_{\text{micro}}(\mathbf{B}) &= \sum_{s=1}^{S} \frac{E(\mathbf{B}_s)}{E(\mathbf{B})} p_s H(\mathbf{B}_s),
\end{aligned}$$

and where p_s is the population proportion of place s among the total metropolitan area population (Mora and Ruiz-Castillo 2011). These two components are referred to as the *macro* (between-place segregation) and *micro* (a weighted sum of within-place segregation scores) components of total segregation. H_{micro} is not an H index, but a weighted sum of H indices. The equation makes clear that this decomposition is simply an accounting scheme: It attributes some part of the total metropolitan area segregation to the "macro" component, and the remainder to the "micro" component. It is not, as Lichter et al. (2015) seem to suggest, a method to discover additional sources of segregation beyond what is already included in $H(\mathbf{B})$ ("Declining segregation at the neighborhood level may be offset by growing segregation between places or other levels of geography, [...] *which are typically excluded from metro-centric segregation studies*", p. 845, emphasis added).

For each combination of metropolitan area and year, Black-White, Hispanic-White, and Asian-White segregation are computed using the H index, each of which can then be decomposed into macro and micro components.

5.4.2 Shapley decomposition

The methodology to decompose changes in segregation using the Shapley value decomposition was described in the previous chapter. The first decomposition is a simple racial-group decomposition, where the changes in segregation are attributed to each racial group's residential mobility patterns. For each metropolitan area, this decomposition is applied separately for the macro and micro components of segregation. In the macro decomposition, the units are places; while in the micro decomposition, the units are blocks. This does not affect the mechanics of the decomposition procedure.

The racial-group decomposition attributes changes in segregation to only two factors. The Shapley decomposition allows us to define arbitrarily complex counterfactuals, with an (in theory) infinite number of factors. In this paper, the interplay between the different spatial units is of special interest. For instance, are segregation dynamics different in suburbs, central cities, and fringe areas? To understand these dynamics, I construct counterfactuals that take into account the type of spatial unit, and whether the racial group in question grew or declined in that unit. Again, I define these decompositions separately for macro and micro decompositions.

For macro segregation, I distinguish between principal cities, suburban places, and fringe areas. By combining these with the racial groups and their growth/decline trajectory, 12 factors of interest are obtained. For instance, the factors for Black-White macro segregation change are: "Black growth in suburban places", "White decline in fringe areas", "White growth in principal cities", etc. The distinction between growing and declining populations is important, because, at least in theory, both components are independent: For instance, the White population could grow in some suburbs (likely increasing segregation), but could decline in some other suburbs (possibly decreasing segregation).

For micro segregation, each block is classified by its racial composition in 1990, following precedent in the literature (Fowler et al. 2016). Any block is classified as "dominant White" if it is more than 90% White, and as "majority White" if it is more than 50%, but less than 90% White. Blocks that are more than 50% Black, Hispanic, or Asian are classified as "majority Black," "majority Hispanic," and "majority Asian," respectively. Because majority Hispanic and majority Asian blocks should not be of much interest when decomposing changes in Black-White segregation, I distinguish only dominant White, majority White and majority Black blocks, while all other blocks are classified as mixed. The equivalent procedure is used for the Hispanic-White and Asian-White decompositions. In total, 16 factors are obtained for each decomposition, with factors such as "Black growth in majority Black blocks", or "White decline in dominant White blocks", etc.

The decompositions are all agnostic to the reasons for "growth" or "decline." Below, I will sometimes equate growth with in-migration and decline with out-migration, but, naturally, population change can occur in the absence of migration due to fertility and mortality. Decomposition factors that account for natural population changes are not incorporated into the decomposition for two reasons. First, such data is not available for small geographic areas such as blocks. One could possibly impute fertility and mortality from aggregated geographies, but these estimates would be very uncertain. Second, in most metropolitan areas, between-neighborhood and between-metropolitan area mobility is much more important quantitatively than fertility and mortality. Note that for fertility and mortality to produce significant changes in segregation, fertility and mortality rates have to differ strongly *between* racial groups. While such differences exists, it is assumed for now that they are not significant drivers of segregation. Winkler and Johnson (2016) show that changes due to natural processes are fairly small compared to changes in net migration. However, their smallest geography is the county, which makes it difficult to compare their numbers to the present study.

5.5 Results

5.5.1 Segregation trends

To assess the evidence that we have for an increasing macro segregation, I start by briefly replicating the results by Lichter et al. (2015). The present research design closely follows their study, but small differences remain: For instance, the present dataset contains 224 metropolitan areas, whereas Lichter et al. (2015) report results for 222 metropolitan areas. Nonetheless, the average segregation indices align very well with those presented by Lichter et al. (2015) in their Table 2. For instance, the correlation for the "total H" between the two studies is 0.99, while the macro and micro terms are correlated at 0.99 and 0.95, respectively. The replication can therefore be considered successful.

Figure 5.1 presents the results of the replication, which shows average segregation across metropolitan areas for each year and racial group comparison. In this figure, average segregation refers to the arithmetic mean across the 224 metro areas. The figure shows some important patterns, most of which are well-known in the literature. First, Black-White segregation is much higher than Hispanic-White and Asian-White segregation. Second, Black-White segregation has seen a large decrease between 1990 and 2010, but still remains at high levels. Hispanic-White and Asian-White segregation have remained at about the same level. Third, in all cases micro segregation is quantitatively more important than macro segregation. Fourth, the macro component of segregation did *not* see a large increase. For Black-White segregation, the macro component has, if anything, slightly decreased (13.4 in 1990, 13.3 in 2010). For Hispanic-White and Asian-White segregation, the macro component did see an increase, albeit from a lower starting point compared to Black-White macro segregation.

Why are these findings seemingly different from Lichter et al. (2015)? The reason is that their paper reports percentages of total segregation, i.e. $100 \times \text{macro}/\text{total}$. This percentage, however,

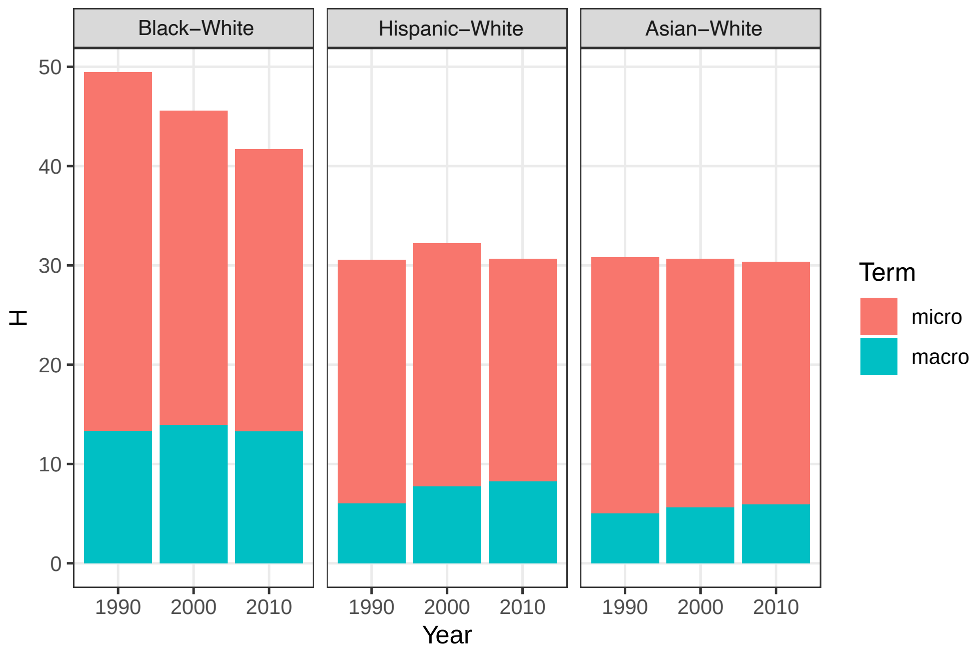

Figure 5.1: Micro and macro segregation, 1990-2010

will increase mechanically if micro segregation declines. Of course, it is true that macro segregation contributes a larger *share* to total segregation in 2010 than it did in 1990, but this is not because macro segregation increased, but because micro segregation declined. Therefore, it is not correct to state that "significantly, macro-segregation—the between-place component—has increased since 1990, offsetting declines in the within-place component" (Lichter et al. 2015, p. 843). Similarly, it is not justified—or at least premature—to state that "[u]rban and suburban municipalities are replacing neighborhoods as the central organising units of metro segregation" (Farrell 2008, p. 467).

Figure 5.1 shows an aggregate picture of residential segregation in the largest metro areas of the U.S., but does not reveal the heterogeneity in trajectories between metropolitan areas. Table 5.2 shows that there is substantial heterogeneity. The table groups metropolitan areas by the size of the change they experienced between 1990 and 2010. For instance, 15 metro areas, or roughly 7%, saw a change in Black-White macro segregation that was higher than 5 on the scale of the H index. For Black-White macro segregation, the metropolitan areas are almost equally split between moderate decline (change of H between -5 and -1), stable segregation (-1 to 1),

Table 5.2: Changes in micro (between-place) and macro (within-place) segregation, 1990 and 2010

		Segregation Change in H					
		< -5	-5 to -1	-1 to 1	1 to 5	> 5	Total
Black-White							
Macro	N	20	56	65	68	15	224
	%	8.9	25	29	30.4	6.7	100%
Micro	N	140	48	15	19	2	224
	%	62.5	21.4	6.7	8.5	0.9	100%
Hispanic-White							
Macro	N	3	26	64	91	40	224
	%	1.3	11.6	28.6	40.6	17.9	100%
Micro	N	56	79	34	41	14	224
	%	25	35.3	15.2	18.3	6.2	100%
Asian-White							
Macro	N	3	32	90	87	12	224
	%	1.3	14.3	40.2	38.8	5.4	100%
Micro	N	34	83	49	53	5	224
	%	15.2	37.1	21.9	23.7	2.2	100%

and moderate increase (1 to 5). However, for micro segregation, 140 out of 224 metro areas saw large decreases. Hispanic-White and Asian-White macro segregation stayed stable or increased moderately in most metropolitan areas (70% and 79%, respectively), while micro segregation for these groups was more spread out, skewing, however, towards declines.

5.5.2 Simple decomposition

The results in Figure 5.1 are based on the cross-sectional files, in order to stay as closely as possible to the design of Lichter et al. (2015). Throughout the remainder of the paper, the results are based on the harmonized file. To assess the sensitivity of the results, I compute the correlation between segregation values obtained using the harmonized and the cross-sectional file across the 224 metropolitan areas. This correlation is 0.999 for Black-White, 0.998 for Hispanic-White, and 0.997 for Asian-White segregation. For micro segregation, the values range from 0.989 to 0.991, while for macro segregation the range is from 0.968 to 0.994. Overall, it can be

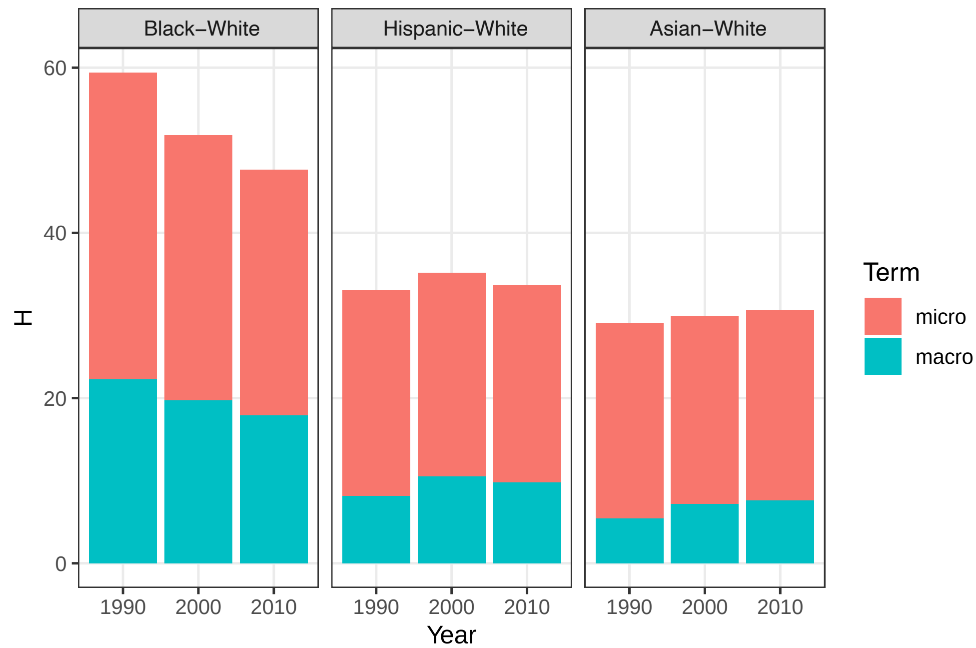

Figure 5.2: Micro and macro segregation, 1990-2010 using the harmonized file

concluded that the use of the harmonized file yields very similar results to the cross-sectional file.

For the decomposition, I also choose a different summary statistic to aggregate results across the metropolitan areas. Lichter et al. (2015) used the arithmetic mean, but this does not take into account differences in population size: Hinesville, GA, with a population of 59,000 in 1990 is weighted as strongly as New York, with a population of over 17 million. I report instead a population-weighted median (see Figure 5.2). For instance, using this measure, the median Black-White segregation in 1990 was around 60, i.e. half of the population lived in metro areas with Black-White segregation above 60, and half lived in metro areas with Black-White segregation below 60. Most patterns are similar between Figures 5.1 and 5.2, but noticeably, using the latter figure we find that Black-White macro segregation actually declined. The question whether macro segregation increased or declined thus also depends on the choice of summary function.

Figure 5.3 presents the results of applying the racial group decomposition to the changes visible in Figure 5.2. The three panels, arranged from top to bottom, show results for Black-White, Hispanic-White, and Asian-White changes in segregation between 1990 and 2010. The figure includes both a population-weighted median for each component in the left-most panel, as well as results for the three largest metropolitan areas of the U.S.: New York–Newark–Jersey City, NY-NJ-PA (population 19.6m in 2010), Los Angeles-Long Beach-Anaheim, CA (12.8m in 2010), and Chicago-Naperville-Elgin, IL-IN-WI (9.5m in 2010). The error bars indicate the population-weighted 25th and 75th percentile, i.e. about 50% of the U.S. population lived in areas where the degree of segregation change fell into the interval shown.

Starting at the top-left, it was already shown that total metropolitan area Black-White segregation declined. The decomposition then consists of four terms, macro change of the White and Black populations, respectively, and micro change of the White and Black populations, respectively. The macro decomposition reveals that, despite macro segregation being essentially unchanged for 20 years, there have been fundamental underlying changes in the population distribution of the Black and White population. In the decomposition, the changes in the population distribution of Whites accounts for changes in segregation of roughly +5. In other words, if the Black population in each place remained stable, but the White population changed according to the observed pattern, average macro segregation would have risen by 5 points—a large effect. The opposite is true of the Black population: changes in the Black population distribution had a desegregating effect. Overall, these two forces effectively offset each other, for an overall small decline in macro segregation.

For micro segregation, the situation is different: Here, both populations contributed to the decline. However, the decline is almost entirely due to changes in the distribution of the Black population. These aggregate patterns match well with the patterns observed in New York and Los Angeles. In Chicago, the patterns for macro segregation are more pronounced: Changes in the distribution of the Black population are associated here with a decline of more than 10 points, a very large decrease. As for micro segregation, both the White and Black population shifts contributed almost equally to the decline. The fact that the White population con-

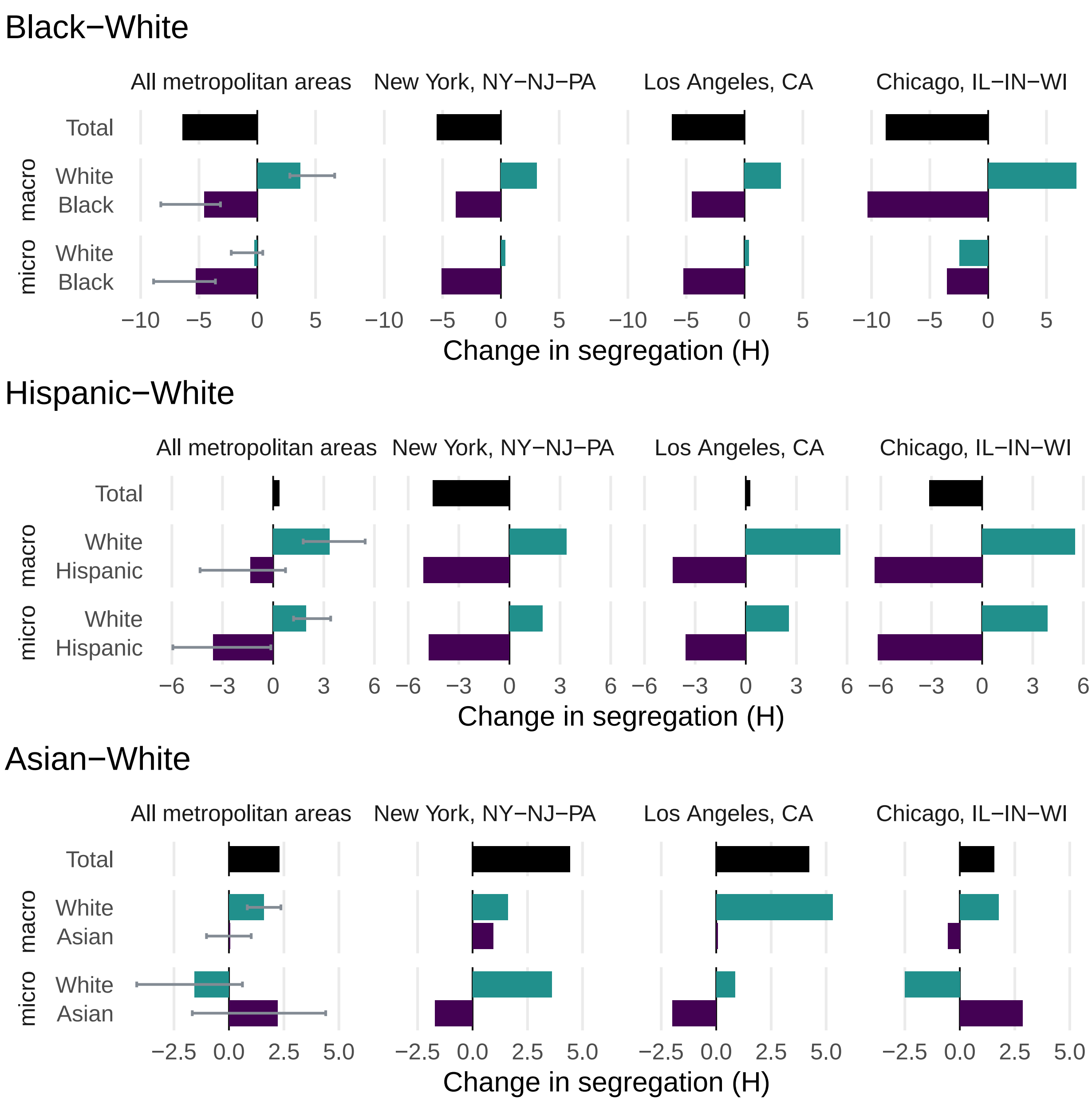

Figure 5.3: Decomposition of changes in segregation, 1990-2010

tributed to decreasing micro segregation in Chicago could be a sign of gentrification, although one would have expected such an effect in the other large metro areas as well.

The decomposition results are similar for changes in Hispanic-White segregation, although the overall magnitude of the changes is smaller. There is one notable difference between Black-White and Hispanic-White segregation patterns: In the latter case, White population shifts contributed towards increasing micro segregation, offsetting a large amount of the decline produced by the shifting Hispanic population. Again, effect sizes are larger in Chicago. Notably, the declines produced by the Hispanic population are large in all three metro areas, leading to pronounced declines in total Hispanic metropolitan area segregation also in New York and Chicago.

Asian-White segregation patterns, shown in the bottom panel, deviate notably from both Black-White and Hispanic-White segregation. The magnitudes of change are again smaller here, ranging from about -2 to 2 points for the average metro area. Notably, while Whites contribute to macro segregation, Asian population shifts in total are not associated with macro segregation changes. While both in terms of macro and micro segregation Blacks and Hispanics reduce segregation, this is not true for Asians: Here, Asian population changes are, on average, associated with increases in micro segregation, although there is large variation across metro areas. Notably, Asian-White segregation has increased by substantial amounts in New York and Los Angeles, mostly due to the increasing micro segregation of Whites in New York, and increasing macro segregation of Whites in Los Angeles.

Figure 5.3 reveals previously hidden patterns of segregation change. In broad strokes, changes in the Black and Hispanic population distributions contributed most to declines in segregation for these groups, while the net contribution of Whites points to increases in segregation. For Black-White segregation, the results are especially stark: Almost the entire decline in segregation is driven by the Black population, and, were it not for the *increasing* macro segregation of Whites, the decline would have been even more pronounced.

A pattern that seems to emerge is that the contributions of Whites run in opposite direction to those of the minority groups. In fact, the correlation between Black and White contributions to macro segregation is large, at -0.6. (For micro segregation, the correlation is -0.16.) The large negative correlation indicates that most metropolitan areas are shaped by simultaneous *de*segregation and *re*segregation. This also explains why macro segregation remained stable: As Blacks and Hispanics moved toward desegregation, Whites continued to move toward segregation.

5.5.3 Detailed macro decomposition

The decompositions in Figure 5.3 show some important patterns of segregation change. However, they do not address the specific spatial sources of change: Is the White contribution towards increasing segregation due to suburbanization? Do Blacks and Hispanics move to the suburbs as well, and thereby decrease segregation? In the detailed decompositions, I decompose macro and micro segregation separately into twelve components each. For macro segregation, the distinction between central city and suburbs is of special interest. For a full picture of metropolitan area spatial arrangements, I distinguish between the Census-defined principal city or cities, suburban places, and fringe areas. Recall that suburban places, in this scheme, only refer to recognized, incorporated entities with some governmental function.

Figure 5.4 shows the decomposition results separately for Black-White, Hispanic-White and Asian-White segregation. As before, these are population-weighted medians, with population-weighted 25th and 75th quantiles shown as error bars. In each panel, the twelve bars sum up to the average change in macro segregation. To interpret the individual factors, consider the example of "Black + in Fringe" for Black-White segregation, where the "+" stands for population growth. To compute this factor, each fringe area is classified in terms of either Black population growth or decline. Then, the counterfactual is computed only for those areas where Black population growth occurred. The factor can then be interpreted as follows: If the Black popula-

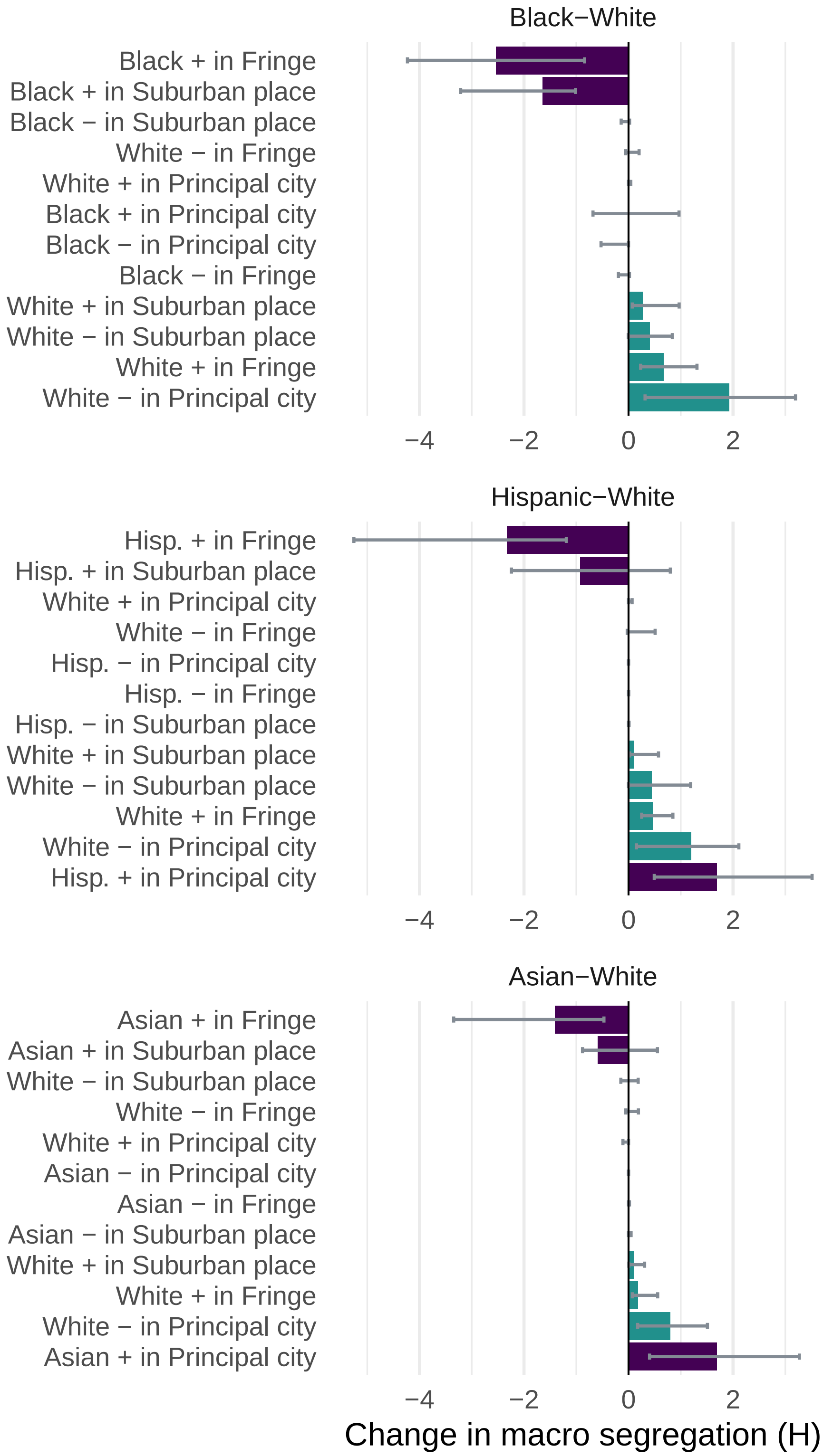

Figure 5.4: Detailed decomposition of changes in macro segregation, 1990-2010

tion grew in those fringe areas where we observed growth between 1990 and 2010, net of other population movements, how would segregation have changed?

The answer is that this factor is large and negative, i.e. −2.5 points of the change in macro segregation is attributed to Black population growth in fringe areas. Another large negative contribution is Black population growth in suburban places. We know that total macro segregation change was basically zero, so it follows that there must be other factors that offset these declines. These are, in order of importance, White suburban growth, White suburban decline, White fringe growth, and, most importantly, White decline in principal cities. Taken in total, these factors are a strong indication of continuing White suburbanization. However, the results also show that there is ongoing sorting *within* suburbs: Note that for Whites, both growth and decline in suburban places contributes towards increasing segregation. This shows that the White population declined in more racially-integrated suburbs, while it increased in more racially-homogeneous suburbs. It is possible that some of these patterns are reactive: Black suburbanization may have prompted Whites to leave suburban places. If this interpretation is true, white flight would not be restricted to principal cities.

It is also interesting to consider effects where we might have expected a contribution, but where the contribution is effectively zero. This is most apparent for the effect of Black decline in principal cities. If there is Black population growth in the suburbs, wouldn't we expect that there is also declining segregation in the principal city? There are several possible explanations for this zero effect. First, it is possible that the fringe and suburban effects reflect a reshuffling of population within suburban and fringe areas. For instance, if Blacks move from majority Black suburbs to majority White suburbs, this reduces segregation in the suburbs, but leaves segregation in the principal city unchanged. Another possible explanation are cross-metropolitan area mobility patterns, where the segregation-reducing effects in the suburbs are brought about by in-mobility from other metropolitan or rural areas.

For Hispanic-White segregation, the overall picture is similar to Black-White segregation. The results show large segregation-reducing effects due to population growth in the fringe and suburbs, and segregation-increasing effects of White suburbanization. An important difference between Black-White and Hispanic-White segregation is that the Hispanic population also contributes to increases in segregation: The largest segregation-increasing factor is Hispanic growth in principal cities. For Asian-White segregation, the effects are smaller, but the patterns are similar. Again, we observe the segregation-reducing effects of Asian suburbanization, as well as the segregation-increasing effects of White suburbanization. As in the Hispanic-White case, a large segregation-increasing effect is due to Asian population growth in principal cities.

These countervailing trends for both Hispanics and Asians suggests that these factors are products of immigration. Newly arrived immigrants often settle in more ethnically homogeneous neighborhoods, where family and friendship networks make for a natural first destination. We can therefore suspect that this group is distinct from the group that suburbanizes: Likely the latter are native-born or later-generation immigrants, who have acquired the resources to move to the suburbs. The Hispanic and Asian patterns are therefore consistent with the spatial assimilation model.

In all three cases, the contribution of minority population growth in fringe areas is greater than in suburban places. There are two possible explanations for this. First, some members of these groups may seek to leave the urban core and move to the suburbs. In traditional suburban places, however, they find barriers that makes moving there either unattractive or impossible. Such barriers are arguably less prevalent in fringe areas, where the additional governmental layer of the incorporated place is missing. This could explain why the desegregating effects are concentrated there. Overall, among those metro areas where both the suburban and fringe components were negative, the fringe component was smaller in 68% of the cases, lending some support to this hypothesis. Another possible explanation is that this effect is explained by aggregating over diverse metropolitan areas. Metropolitan areas in the fast-growing South and West of the U.S. have a lower density of places, and a larger share of the population in fringe areas. (For instance, in 2010, 19% of the population of Houston resided in suburban places, but 41%

in fringe areas. In Chicago, the numbers are 56% and 8% respectively.) If segregation declines are larger in those areas, this would explain why the fringe component is more important.

5.5.4 Detailed micro decomposition

The results for the detailed decomposition of micro segregation change are shown in Figure 5.5. As described above, this decomposition uses blocks as the spatial units. The decomposition is carried out separately for principal cities, suburban places, and fringe areas. The rationale for this is that we have seen large differences between these areas for macro segregation change, and we may therefore expect that these areas also differ fundamentally for micro segregation change.

Focusing first on Black-White segregation, Figure 5.5 shows that the largest factor affecting micro segregation across all three geographical subdivisions is Black population growth in dominant White areas. One might have suspected that this is mostly a suburban and fringe effect, but the effect is large also for principal cities. The effect is smallest in suburban places, which might point to the fact that place-based barriers hinder in-mobility of minorities. A second important factor, especially in principal cities, is Black population growth in majority White blocks. Note that a Black family moving to a dominant White area (90%+ White) brings about a much larger decline in segregation than the same family moving to a majority White (50-90% White) area. The results are consistent with this pattern.

The decline in micro segregation that we observe is therefore mostly due to Blacks moving into areas that were more than 50% White in 1990, and this is the case in all parts of the metropolitan area. For principal cities, we also observe a large segregation-reducing effect for Black population decline in majority Black areas. This effect is smaller in suburban and fringe areas, likely because majority Black areas are rare outside principal cities.

Factors that increase segregation are smaller, but there are quite a few population movements that partially offset the overall decline. These are the classic indicators of white flight (White population declines in majority Black and mixed neighborhoods). One factor that in-

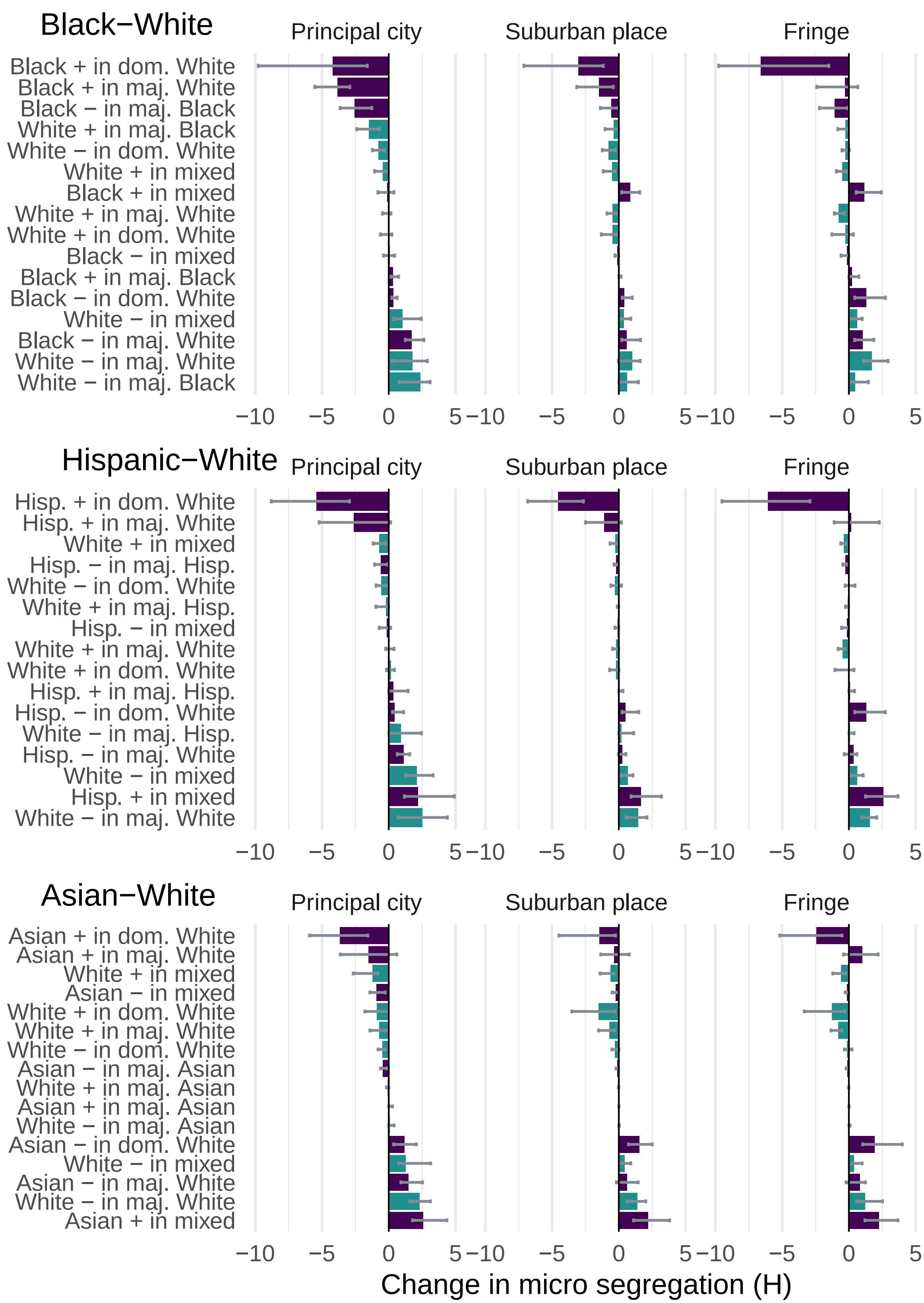

Figure 5.5: Detailed decomposition of changes in micro segregation, 1990-2010

creases segregation is also White population decline in majority White areas, which may be surprising. However, given that many Whites are segregated in neighborhoods that are more than 90% White, the White population in neighborhoods that are "only" 50-90% White is in fact relatively integrated. If Whites move out of these neighborhoods, the overall effect is therefore segregation-increasing. Blacks moving out of majority and dominant White areas also contributed to the offsetting factors, showing again that there is substantial heterogeneity across neighborhoods. While some majority White neighborhoods experience minority population growth, others experience a decline. Similarly, in suburban places and fringe areas, we observe a segregation-increasing effect of Black population growth in mixed blocks, possibly an indicator that subsets of suburban and fringe areas are becoming more racially homogeneous.

Nonetheless, the main driver of changes in Black-White segregation is minority population growth in majority White areas. The total effects, summing over all twelve factors, are -5.8 for principal cities, -3.9 for suburban places, and -3.6 for fringe areas. The pattern is similar for Hispanic-White segregation, with a similarly large effect of Hispanic population growth in dominant and majority White areas. However, the total effects here are only -0.5 (principal cities), -2.2 (suburbs) and -0.6 (fringe). This explains why micro segregation declines for Hispanics have been lower than for Blacks: Although these groups have both moved to majority White areas, the offsetting (segregation-increasing) factors are larger for Hispanics. In principal cities, one large offsetting factor is the decline of the White population in mixed, and to a lesser extent, majority Hispanic areas ("white flight"). Another significant factor is Hispanic population growth in mixed areas. The fact that this movement is associated with segregation increases strongly suggests that the "mixed" areas are usually neighborhoods where Whites are significantly underrepresented. Again, Hispanic mobility towards majority White and mixed neighborhoods may also be due to internal heterogeneities within the Hispanic group: Immigrants may settle predominantly in majority Hispanic and mixed neighborhoods, while established members may have acquired the resources to move into majority White neighborhoods.

Only for the Asian-White decomposition do we see larger differences between principal cities on the one hand, and suburban places and fringe areas on the other hand. The total effects here are -1, 0.9, and 2.1. Thus, micro segregation declined slightly in principal cities, increased slightly in the suburbs, and increased by a large amount in fringe areas. The largest segregation-decreasing effect is, again, Asian growth in dominant White blocks. Unlike for Blacks and Hispanics, we observe large segregation-decreasing effects of Asian population growth in majority White blocks only in principal cities. The set of segregation-increasing factors is diverse, and similar to the other two comparisons. For Asians, the heterogeneity of effects is especially visible: In suburbs and fringe areas, the segregation-reducing effect of Asian *growth* in dominant White areas is entirely offset by the segregation-increasing effect of Asian *decline* in dominant White areas. This simultaneity of both positive and negative developments is what characterizes changes in micro segregation for Asians and Hispanics.

Lastly, especially for Hispanics and Asians, the "mixed" blocks play a large role in producing segregation-increasing effects. In accordance with the spatial assimilation model, this may be because these neighborhoods serve effectively as stepping stones between majority Hispanic or Asian areas, and higher-status suburban and fringe areas. It is also possible that we observe these neighborhoods in a transitory stage, as Bader and Warkentien (2016) have argued.

5.6 Discussion and conclusion

Despite the complexities of the individual population movements, there are clear patterns that characterize segregation change in U.S. metropolitan areas between 1990 and 2010. One significant dimension is the distinction between macro and micro segregation. I find, consistent with earlier research, evidence that macro and micro segregation are distinct processes. The clearest example of this are the contributions of different racial groups to segregation change. Especially for Black-White segregation, macro segregation is a product of both White and "minority" population movements. Minority population movements are typically segregation-reducing, while White population movements are typically segregation-increasing.

In total, these countervailing trends produce a trend in macro segregation that is stable (for Black-White) or slightly increasing (for Hispanic-White and Asian-White segregation). These offsetting trends show that one should not mistake the stability of the aggregate measure for the absence of change.

The stability of macro segregation contrasts with declining Black-White micro segregation. This is the most significant trend in the data: Using the harmonized file, average Black-White segregation declined from 37 to 30 points, a decrease of 19%. Why did Black-White micro segregation decline by such a large amount, while macro segregation remained stable? The reason is that there was no significant countervailing movement of Whites, at least for Black-White segregation. Blacks and Hispanics moved towards integration, and Whites did not, in large numbers, move toward segregation. In fact, the detailed decompositions show that White population movements overall matter very little for micro segregation; the largest effects are all produced by the minority groups.

The stability of macro segregation offers support for the place stratification hypothesis. This is also supported by other evidence, such as the fact that the contribution of minority population growth in (unincorporated) fringe areas is greater than in (incorporated) suburban places. Places are clearly important in the production and maintenance of segregation, but, as of 2010, micro segregation is still quantitatively more important than macro segregation. If the trend of declining micro segregation continues, this may soon change. An open question in the literature is also why racial residential macro segregation is stable, while poverty and income segregation have risen (Owens et al. 2016; Reardon et al. 2018). Future research should aim to explain this mismatch between racial- and class-based forms of inequality.

The results of this study also show that models such as place stratification or spatial assimilation may not be very useful in characterizing aggregate changes in metro area segregation. Especially the detailed decompositions reveal that there are several important population movements that are consistent with *either* model. For instance, as Asians move to dominant White areas, consistent with spatial assimilation, other dominant White areas experience a decline in the Asian population, consistent with place stratification. The overall picture that segregation change presents here, it seems, is best characterized through the "simultaneity of processes." Desegregation and resegregation have happened, and are likely still happening, concurrently.

Clearly, to fully explain all the population movements that have been identified as significant in either increasing or decreasing segregation throughout this study, future research and more detailed data is needed. For one, the current study does not distinguish population movements that are brought about by natural population changes (fertility and mortality) from migration. For a deeper understanding of the migration component, it would also be desirable to distinguish between short-distance migration within the metropolitan area, long-distance migration across metropolitan areas, and immigration. Such data is hard to obtain, especially for smaller geographies such as blocks.

Another possibility of segregation change is the declining significance of racial identification. The meaning of race in the U.S. may have undergone significant changes since 1990. For instance, changing rates of marital homogamy and the possibility to identify as multiracial may weaken the strong majority-minority dichotomy in American race relations. Contrary to the assertion that populations need to grow or decline to change segregation, if individuals regulary change their racial identification, this process can impact segregation without population mobility. It would be useful to incorporate these dimensions of segregation change into future research.

Chapter 6: Recent Changes in School Segregation

6.1 Introduction

Given the consequences of racial segregation for individual life outcomes and its role in exacerbating inequalities in the U.S. (Reardon and Owens 2014), it is of major importance to better understand the processes that shape segregation. While we know a lot about national trends in racial school segregation, the factors that have shaped these trends since the 1980s are little understood. Case studies focus on *increases* in school segregation: results show an expansion of the "school choice" movement (Saporito 2003), the halting of desegregation efforts in the South (Orfield et al. 2002; Reardon et al. 2012), and increasing economic segregation in schools (Owens 2016). However, studies at the national level find that the racial segregation of schools has either stayed stable or even declined since the late 1980s (Stroub and Richards 2013). These results imply that there are countervailing factors that decrease school segregation, such as declines in residential segregation (see previous chapter). To advance our understanding of why school segregation is still so prevalent in the U.S., but also why it is slowly declining, the analysis must shift from the *description* of trends to the *explanation* of trends. The approach taken in this chapter is to more immediately connect changes in school segregation to relevant demographic trends. How much does the residential mobility of each racial group contribute to increases and decreases in school segregation? And what are the major countervailing forces that stabilize school racial segregation?

Today, over two-thirds of school segregation in U.S. metropolitan areas is between school districts, while about one-third of school segregation occurs within districts (Stroub and Richards 2013). However, the between-district component has generally received less attention in the literature. One reason for this may be that most school desegregation efforts targeted within-district segregation. Busing, for instance, occurred within school districts. In the South, school districts are generally large, and the focus on within-district segregation seemed reasonable. However, it was soon observed that efforts aimed at desegregating school districts often led to increases in between-district segregation, especially between cities and suburbs (Coleman et al. 1975). This problem is aggravated in the cities of the Northeast and Midwest, where school districts are often small, and where there is hence much more opportunity for differentiation between school districts. As between-district segregation is so large, and in some parts of the country even increasing, I focus on this aspect of school segregation in this paper.

Focusing on the between-district, within metropolitan area component of school segregation also implies a wider perspective shift. While trends in school segregation are often described using the metropolitan area as the basic unit of analysis (e.g., Owens 2018), studies of achievement gaps between racial groups often take the school district as the basic unit of analysis (e.g., Reardon et al. 2019b). Such studies find large differences between states and regions in terms of average achievement, and in terms of achievement gaps between racial groups. The focus on metropolitan areas as the basic unit of analysis (as in this paper) implies that relative differences within metro areas are equally important. Estimates from the Current Population Survey (CPS) for the period 2010-2020 show that about 60-70% of all moves are within the same county, while only 13-15% of moves cross state lines.[1] This implies that many parents, when evaluating potential residential locations, will focus on locations within the metropolitan area, or mostly consider even shorter distance. (Clark (1986) explicitly argues that the fact that most moves happen over short distances is one of the main reasons for the stability of segregation.) Hence, for residential mobility, the differences between school districts will likely be what parents focus on, although another metropolitan area may offer better overall schools.

Compared to the previous chapter, this chapter focuses on school segregation, not residential segregation. While school and residential segregation are related, they are not identical, as

[1] https://www.census.gov/data/tables/time-series/demo/geographic-mobility/historic.html

the student population is younger, and therefore recent immigration and fertility changes are more immediately present. This chapter also focuses on more recent changes in school segregation (2009-2016 for the descriptive results, 2010-2015 for the decomposition results), for which there is more detailed data available than for earlier periods. Lastly, the fact that students advance regularly by one grade from year to year makes it possible to control for changes in fertility. If there are $n_{g,t=1}$ students in grade g at time $t = 1$, we would expect a similar number of students in grade $g+1$ at time $t = 2$. Under plausible assumptions, the difference in student counts, $n_{g+1,t=2} - n_{g,t=1}$ can then be attributed mostly to residential mobility. This is the key fact that is exploited in the research design for this paper, where the interest lies in connecting selective changes in residential mobility to changes in school segregation.

For this study, I use the Stanford Education Data Archive (SEDA, Reardon et al. 2019a), which combines National Center for Education Statistics (NCED) enrollment data for public schools, nationwide test score data obtained from the Department of Education, and American Community Survey (ACS) data. This dataset allows me to estimate detailed decompositions that connect residential mobility flows to changes in school segregation. I argue that, once we know what factors explain trends in school segregation, more effective policy responses can be designed. Policy responses need to focus on preventing increases in school segregation, but also focus on strengthening the countervailing forces that reduce segregation—a fact that has not received much attention in the literature.

6.2 Background

The main policy and research focus with respect to school segregation has been on within-district segregation. Within-district segregation refers to that part of total segregation in a metropolitan area that arises from the unequal distribution of racial groups across schools within the district's purview. The amount of within-district segregation is based on the assignment procedures that each school district uses to allocate students to specific schools. Within-district segregation was thus a natural starting point for desegregation plans.

Court-ordered desegregation plans first targeted school districts that had separate schools for Black and White students, and were then later extended to highly segregated districts (Coleman et al. 1975). As Logan and Oakley (2004) show, the effects of desegregation spread to school districts that were under no court order as well, possibly to preemptively avoid a court order. These effects led to rapid declines in Black-White segregation within school districts up until the 1980s (Welch and Light 1987). Briefly, it was also considered to extend desegregation to exchanges of students between school districts, but since the 1974 Supreme court decision *Milliken v. Bradley* set high requirements for inter-district busing, this was attempted only in rare cases.

However, it quickly became apparent that school desegregation, achieved mainly through the integration of dual school systems and busing, may have had an unintended consequence in the form of "white flight." In 1966, the Coleman Report found that segregation explained an important part of the Black-White achievement gap and thus argued in favor of desegregation (Coleman et al. 1966); by 1975, however, Coleman and colleagues argued that within-district desegregation was harmful because it led to an increase in between-district segregation (Coleman et al. 1975). This foundational finding has sparked a series of studies that, overall, produce conflicting results. Responding directly to Coleman, Rossell in 1975 found no relationship between desegregation efforts and white flight (see also Clotfelter 2001; Farley et al. 1980; Frey 1979; Smock and Wilson 1991). In a later study using a different sample and time frame, Rossell (1978) concluded that desegregation triggers white flight, but that the net effect of desegregation plans remains positive. Reber (2005) found weak evidence that desegregation is followed by white flight, but found that a higher number of school districts in an area increased white flight. Reardon and Yun (2001) show that white flight is not limited to moves from the central city to the suburbs, but that in-migration of minorities into suburbs increases suburban segregation as well.

Another form of white flight is private schooling, a topic of intense debate in the early 1980s (Coleman et al. 1982a,b; Taeuber and James 1983). Recent studies find that private schools contribute to racial segregation in schools, although overall enrollment rates are stable (albeit with differences by race and family income, see Murnane and Reardon 2018). Andrews (2002) finds that desegregation plans are associated with the establishment of private academies for White students in the South. For 1999–2000, Clotfelter (2011) estimates that about 16% of segregation in metropolitan areas is due to private school enrollment. Recently, other forms of school choice have been investigated, such as the segregative effects of charter schools. Renzulli and Evans (2005) show that when school districts are integrated, White charter school enrollment increases. Bifulco and Ladd (2007) and Ladd et al. (2015), focusing on North Carolina, confirm these results, and also show that charter schools contribute to the Black-White achievement gap (see also Frankenberg et al. 2010).

Despite these findings that point to increases in segregation—produced by white flight, private schools, and charter schools—, studies that document nationwide trends in school segregation find that school segregation has remained at roughly the same level or even decreased since the 1980s. For the period 1989–1995, Reardon et al. (2000) find a 1% increase in multiracial segregation, and document a decrease in within-district segregation (partially offset by larger increases between districts) (see also An and Gamoran 2009). Logan et al. (2008) study a longer period (1970–2000) and find a similar pattern of change between 1990 and 2000. Stroub and Richards (2013) study trends for 1993–2009 and find that multiracial segregation has decreased by 10%, with declines in both the within- and the between-district component.

The findings of these studies rely on "evenness" measures of segregation, such as the index of dissimilarity or the *H* index. Other studies have used exposure measures (measuring the share of White students that the average Black student encounters in school). If exposure measures are used, studies find that segregation has increased strongly (Frankenberg et al. 2003; Orfield and Yun 1999). However, many studies have shown that exposure measures conflate demographic and segregation change (An and Gamoran 2009; Fiel 2013; Logan 2004), and thus cannot be used to reliably detect segregation trends. Fiel (2013) explicitly decomposed the exposure index over time and found that compositional changes account for the declining presence of White students in minority schools. The consensus in the sociological literature has now built around stable or declining rates of school segregation in recent years (Reardon and Owens 2014), but there remains considerable disagreement in the media.[2] The question of whether segregation increased or decreased should not distract, however, from the fact that the absolute levels of racial segregation in U.S. schools remain very high.

Currently, the literature has not resolved the conflict between the evidence that documents increasing segregation, often established through case studies, and the stable patterns of segregation that are shown in nationwide trend studies. Despite the increasing importance of the between-district component, many studies continue to focus exclusively on within-district segregation, such as the segregative effects of school attendance boundaries (Richards 2014; Saporito and Sohoni 2007; Saporito 2017a,b). Similarly, trend studies rarely discuss the between-district component in depth. Logan et al. (2008) interpret the increase in the between-district component as evidence for white flight, but do not explicitly test this proposition—in fact, their models are inconsistent with this hypothesis, as shown below. This also raises the question why white flight should continue if desegregation plans are no longer in place.

Some researchers use regression models at the level of the metropolitan area or the school district to study the effects of population and minority growth, international immigration, or private schools on segregation levels. Logan et al. (2008), in cross-sectional models for the years 1970 to 2000, find that desegregation orders, the percentage of children in private schools, and whether school district boundaries comprise both the city and suburbs[3] are factors associated with lower metropolitan area segregation. Residential segregation, larger school districts, higher income differences between Blacks and Whites, and the percentage of Black children are factors associated with higher segregation. These results are roughly similar in models that separately estimate the same model for within- and between-district segregation. One exception is the effect of desegregation orders, which is positive for within-district segregation, but negative

[2]For instance, see this post on Vox: "The predominant narrative among education activists is that school segregation has gotten worse in the past several decades. It's an argument backed by data." (`vox.com/2018/3/5/17080218/school-segregation-getting-worse-data`).

[3]Effectively an indicator of school district size.

for between-district segregation. If anything, one would have expected the exact opposite pattern. The authors explain this finding through reverse causation (especially segregated districts are targeted first for desegregation orders), but the difficulty of explaining many of the regression coefficients clearly shows that metropolitan-area, cross-sectional models are not ideal for explaining localized processes of segregation.

A better understanding of the between-district component of school segregation requires knowledge about the underlying demographic processes, which have been only partly considered until now. Owens (2017) matched school district boundaries to the underlying Census geographies and found that children are more segregated than adults. This suggests that families strategically choose their residence with respect to school districts (see also Brunner 2014), and could potentially also explain divergent trends between school segregation and residential segregation. Fiel and Zhang (2018) decompose school segregation using an Age-Period-Cohort (APC) model. Consistent with earlier evidence, they find that there is a declining period effect between 1999 and 2013, but they unexpectedly find an increase in the cohort effect. This shows that the further decomposition of segregation changes into a number of (possibly offsetting) components is useful.

6.3 Data

I use the Stanford Education Data Archive (SEDA, Reardon et al. 2019a), which combines National Center for Education Statistics (NCED) enrollment data for public schools, nationwide test score data obtained from U.S. Department of Education, and American Community Survey (ACS) data. Most importantly, SEDA places all test scores on a common scale. This is necessary because different U.S. states administer different tests, and these tests are also not stable over time. SEDA also accounts for school district boundary changes over time and merges covariates from the ACS to the school districts. Currently, SEDA contains data for students in grades 3 through 8 for the period 2009-2016. For this paper, I make use of the average yearly test scores by school districts, although the data contains much finer detail (e.g., test scores by subgroups and grade levels).

SEDA contains only metropolitan divisions, many of which include only one school district. To improve comparability with the previous chapter, I merge the relevant metropolitan divisions to arrive at metropolitan areas. I use the 200 largest metropolitan areas by enrollment, and use official Census categories to classify each metropolitan area by region (Midwest, Northeast, South, and West). I remove seven metropolitan areas that still contain only one school district. I remove these because their between-district segregation scores will be zero by definition. The remaining 193 metro areas represent about 75% of the total enrollment of public school students in the U.S. in grades 3-8 in a given year.

As the SEDA data is based on administrative data, it is generally of high quality. Some data entry errors occur, however. For instance, the Elk Grove Unified school district in the Sacramento, CA metropolitan area reports 4,743 students in grade 5 in 2009, 4,793 students in grade 6 in 2010, 350 students in grade 7 in 2011, and 4,887 students in grade 8 in 2012. Based on the grade progression, the reporting of 350 students in 2011 is likely the result of a data entry error when the state or school district submitted the results to NCES. I use a simple heuristic to fix such entries: Outliers are identified through sudden breaks in grade progression, defined as a change that exceeds 25% increase or decline in grade progression year-over-year, where the enrollment is at least 100 in either year. Only 816 cases (0.3% of year-grade-district combinations) are identified through this method. Therefore, I assume that the result of data entry errors is negligible. Nonetheless, I replace the outliers for these cases with the prediction from a linear model that interpolates the grade progression for the specific school district.

The full data is described in Table 6.1, where the number of metro areas, districts, and students, as well as their racial composition, are broken down by Census region: Midwest, Northeast, South, and West. As will be seen below, average segregation varies widely across these regions. In terms of racial composition, there are also a number of clear patterns: While White students still are in the majority in the Northeast and Midwest, their numbers have declined by between two and six percentage points in all regions. The percentage of Black students has declined slightly, and is highest in the South. The Hispanic population has grown everywhere by between four and six percentage points, and has reached almost 50% in the West. The Asian student population has stayed rather stable over this period.

For the decompositions, I will make use of the cohort structure of school enrollment data. For instance, a child that was in grade 3 in 2010, would under normal circumstances be in grade 4 in 2011, in grade 5 in 2012, and so on. A child in grade 3 is usually 8 years old, so I refer to this example cohort as the 2002 birth cohort. Table 6.3 shows the Lexis table of grade and year progression, marking the three birth cohorts 2001, 2002, and 2003, for which the maximum number of observations is available in the data. The 2002 birth cohort dataset would then refer to the subset of students who belong to the 2002 birth cohort. The key idea exploited here is that, for a given birth cohort, fertility effects are controlled for. While students can exit the birth cohort due to mortality, it is unlikely that this will affect the estimates. Differences in the racial composition between a given birth cohort in grade 3 and the same birth cohort in grade 8 can then be plausibly attributed to residential mobility, including immigration from abroad.

One problem for the research design of this paper is when students mechanically switch school districts at some grade transition. This will be the case in areas where elementary and middle schools are served by different school districts. For instance, in California, there will often be an "Elementary School District" and a "High School District" (which also serves grades 7-8) in a given area, with overlapping geographies. In a "unified" school district, which serves all grades, we would expect that about one third of the students are found in grades 7 and 8, as the data contain six grades in total. I mark only those districts as "unified" where between 20% and 40% of students are found in grades 7 and 8, otherwise the district is marked as "non-unified." These non-unified districts are included in the descriptive results, but will be dropped for the decomposition analyses. In 2010, 369 out of 5737 school districts (6.4%), representing about 1.5% of the student population, are non-unified.

Table 6.1: Summary statistics across Census regions

	N			Racial Distribution (in %)			
Year	Metro	Districts	Students*	White	Black	Hispanic	Asian
Midwest							
2009	40	1,683	3,179	65	20	11	4
2010	40	1,715	3,179	65	19	12	4
2011	40	1,716	3,172	64	19	13	4
2012	40	1,716	3,170	63	19	13	4
2013	40	1,723	3,171	63	19	14	4
2014	40	1,722	3,169	62	18	14	4
2015	40	1,723	3,163	62	18	15	4
2016	40	1,693	3,167	61	18	15	5
Northeast							
2009	28	1,794	3,088	60	17	17	7
2010	28	1,796	3,084	59	16	17	7
2011	28	1,800	3,125	58	17	18	7
2012	28	1,795	3,110	57	16	19	7
2013	28	1,797	3,110	56	16	20	7
2014	28	1,796	3,100	55	16	21	8
2015	28	1,815	3,095	54	16	22	8
2016	28	1,818	3,099	54	16	22	8
South							
2009	79	1,099	5,933	45	27	24	4
2010	79	1,101	6,012	44	27	25	4
2011	79	1,095	6,075	43	26	26	4
2012	79	1,096	6,136	43	26	27	4
2013	79	1,103	6,178	42	25	28	4
2014	79	1,100	6,234	42	25	28	4
2015	79	1,102	6,302	41	25	29	4
2016	79	1,104	6,391	40	25	30	4
West							
2009	46	1,131	4,353	39	7	43	10
2010	46	1,125	4,322	39	6	44	10
2011	46	1,133	4,346	38	6	45	9
2012	46	1,127	4,359	38	6	46	9
2013	46	1,127	4,383	38	6	46	9
2014	46	1,127	4,412	37	6	47	9
2015	46	1,128	4,451	37	6	47	9
2016	46	1,126	4,482	37	5	47	9

* in 1,000

Table 6.2: Cohorts available in the SEDA data

Year \ Grade	3	4	5	6	7	8
2009	2001	2000	1999	1998	1997	1996
2010	2002	2001	2000	1999	1998	1997
2011	2003	2002	2001	2000	1999	1998
2012	2004	2003	2002	2001	2000	1999
2013	2005	2004	2003	2002	2001	2000
2014	2006	2005	2004	2003	2002	2001
2015	2007	2006	2005	2004	2003	2002
2016	2008	2007	2006	2005	2004	2003

Another potential confounder of mobility patterns could be private schooling, and school (district) choice. Given that I only study between-district segregation, school choice is less of a problem in this setting. Even if students switch schools (say, from primary school to middle school), this school will in most cases be in the same school district, as I focus on unified districts. Private schooling is possibly a bigger issue, if a large number of students moves in or out of private schools after grade 3 and/or before grade 8. I am not aware of datasets that provide information on switchers between private and public schools (ideally broken down by grade), and I will assume for now that most children are educated in either the public or private school system for the majority of their schooling careers. For private school switching to seriously affect the estimates, the *switching* (not absolute differences in private school attendance) would also have to vary greatly between racial groups, and the *net* change has to be either positive or negative. Hence, under these assumptions, changes in the racial composition that are observed by following a birth cohort in a specific school district over time can be attributed entirely due to selective in- and outmigration.

6.4 Methods

As before, I use the entropy-based segregation index H. For a given metropolitan area, let $\mathbf{T}$ be a matrix with U rows (here, school districts) and G columns (racial groups). Let the entries of $\mathbf{T}$ be t_{ug}, the number of people of race g in school district u, and let t be the total population of $\mathbf{T}$, i.e. $t = \sum_{u=1}^{U} \sum_{g=1}^{G} t_{ug}$. The joint probability of being in school district u and racial group g is $p_{ug} = t_{ug}/t$. Also define $p_{u\cdot} = \sum_{g=1}^{G} t_{ug}/t$ and $p_{\cdot g} = \sum_{u=1}^{U} t_{ug}/t$ as the marginal probabilities of school districts and racial groups, respectively. The H index is then defined as

$$H(\mathbf{T}) = \frac{100}{E(\mathbf{T})} \sum_{u} \sum_{g} p_{ug} \log \frac{p_{ug}}{p_{u\cdot} p_{\cdot g}},$$

where $E(\mathbf{T}) = -\sum_{g=1}^{G} p_{\cdot g} \log p_{\cdot g}$ is the entropy of the racial group marginal distribution of $\mathbf{T}$. In this formulation, the index ranges from 0 (absence of segregation) to 100 (complete segregation).

For each combination of metropolitan area and year, I calculate Black-White, Hispanic-White, and Asian-White segregation using the H index. Below, I also compute H indices for subsets of specific birth cohorts, and I compute the multigroup H index that involves the segregation between the White, Black, Hispanic, and Asian student populations simultaneously.

Similarly to the last chapter, I decompose the change in the H index into a number of factors, using the Shapley decomposition described in chapter 4. However, in this chapter, I focus on the period between 2010 and 2015 and on the 2002 birth cohort to control for segregation change that is due to cohort effects. I compute three separate decompositions. The first decomposition attributes changes in segregation to the residential mobility of specific racial groups. This is a simple decomposition involving just two factors. The next set of decompositions, as before, is based on the racial composition of the school district in 2010. The last set of decomposition is based on the average *student performance* of the school district, which is operationalized here as the average test score achieved in the school district.

6.5 Results

6.5.1 Segregation trends

Figure 6.1 shows the trends in between-district school segregation, separately by Census regions. All estimates are weighted by the student population size of the metropolitan area. The first panel shows the multigroup H index, based on the four racial groups. In the Midwest and Northeast, multigroup segregation declined by about -3. In the South and West, segregation is much lower, and has declined very little.

Black-White, Hispanic-White, and Asian-White segregation, shown in the three additional panels, exhibit different patterns, justifying their separate treatment. Black-White segregation in the Midwest and Northeast—which have the highest segregation across all indices and regions—declined by about -2, while segregation stayed stable in the South and West. Hispanic-White segregation is highest in the Northeast, where it also has declined the most (-5). Hispanic-White segregation has also declined in the Midwest and West, while it has stayed stable in the South. Asian-White segregation has increased everywhere, but especially in the Midwest. In absolute terms, high levels of between-district segregation are only observed for Black-White segregation in the Midwest and Northeast, and for Hispanic-White segregation in the Northeast. It is, however, worrying that segregation has increased for Asians, and is not declining strongly in other areas of the U.S.

6.5.2 School district fragmentation

Between-district segregation arises from the differential distribution of racial groups across school districts. This implies that the potential for segregation is greater when the number of school districts within a metropolitan area is large. Indeed, the number of school district varies greatly between metropolitan areas. For instance, the Boulder, CO metropolitan area comprises only two school districts, while the New York City metropolitan area contains over 500. (However, many of these are special school districts and very small.) The number of school districts

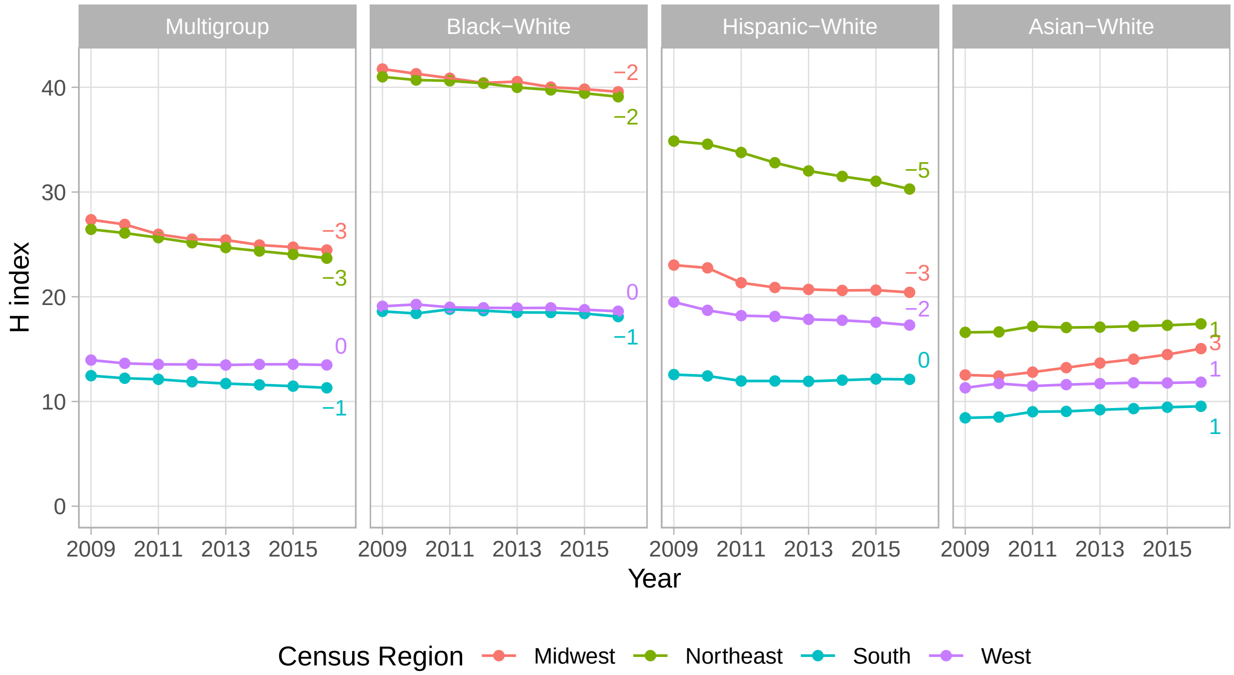

Figure 6.1: Trends in between-district school segregation
Note: Numbers indicate change in the H index between 2009 and 2016.

per area is often used as a measure of school district fragmentation. However, this measure does not incorporate the distribution of students over school districts. If the metropolitan area has many districts, but most of these are very small, the potential for segregation is smaller than when the students are approximately equally distributed across districts. Hence, a better measure of district complexity is the entropy of the distribution of students across school districts, i.e.

$$E(\mathbf{p}_{u\cdot}) = -\sum_{u=1}^{U} p_{u\cdot} \log p_{u\cdot},$$

where $p_{u\cdot}$ is the proportion of the student population of district u among the total enrollment of the metropolitan area. The entropy grows with an increasing number of school districts, but grows especially strongly when students are equally distributed across school districts. Therefore, the entropy incorporates the two necessary dimensions that capture the "potential" for school segregation. A similar measure of fragmentation was used by Bischoff (2008):

$$F(\mathbf{p}_{u\cdot}) = \sum_{u=1}^{U} p_{u\cdot}(1 - p_{u\cdot}).$$

When all students are concentrated in one district, this measure is zero. When there is a large number of districts, each of which is approximately of equal size, the measure approaches 1.

Figure 6.2 shows the entropy of the school district distribution compared to the segregation index for the year 2010. The plot shows a positive correlation, which is about 0.73. For comparison, the correlation between Bischoff's fragmentation index and the H index is 0.65, somewhat less predictive. (The fragmentation index and the entropy correlate at 0.9 in this dataset.) The correlation between the number of districts and segregation is only 0.4, hence, either the fragmentation index or the entropy is preferred for investigating school district fragmentation.

Despite the strong positive correlation visible in Figure 6.2, it is also clear that the relationship is mediated by geography: Large cities in the Midwest and Northeast are located in the upper-right area of the plot, while most cities of the South and West are located in the bottom-left. This suggests that one reason for the lower levels of segregation in Southern and Western U.S. metropolitan areas is the relative lack of school district fragmentation.

The argument that school district fragmentation is a strong driver of segregation is related to the idea of macro segregation. School districts are similar to "places" in that they are independent jurisdictions, differing widely in the resources that are available to them. Studies have found that it is highly plausible that school districts play an important part in guiding the parents' choice of residential location (e.g., Bischoff 2008; Brunner 2014; Owens 2017). Bischoff (2008) specifically observes that fragmentation is unrelated, or even negatively related, to within-district segregation. This could suggest that parents are primarily concerned with school district boundaries, and less concerned with the exact residential location within a school district.

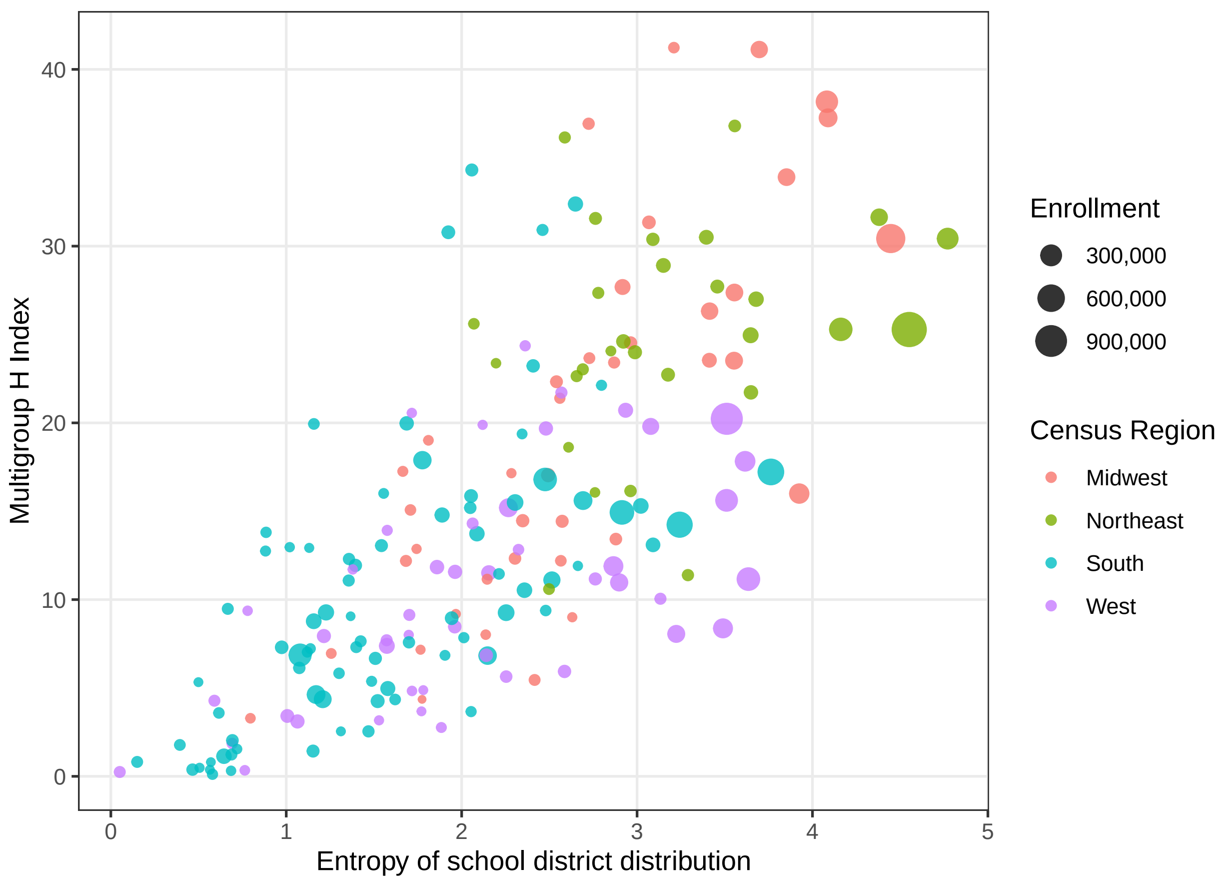

Figure 6.2: Multigroup segregation and entropy of the school district distribution

6.5.3 Test scores

If school district choice is one of the important dimensions on which at least some parents select a residential location, the question is on which district characteristics the choice is based. Especially in highly fragmented metropolitan areas—Bischoff (2008, p. 208) cites the case of Bergen County, New Jersey, where 74 school districts encompass an average of less than four square miles—, school districts will vary widely on characteristics such as housing prices and availability, access to transportation, and distance to jobs, all of which are important for residential location decisions (Cadwallader 1992). However, school districts also vary by student performance, which may be an important observable indicator for parents. The SEDA data contains standardized test scores comparable across states, which makes it possible to use test scores as an indicator of school district performance.

Figure 6.3 shows the variation in median test scores across metropolitan areas by selecting the top and bottom 10 metropolitan areas according to the test score distribution. The quantiles shown are population-weighted across the school districts' average test scores. For instance, in the Bridgeport-Stamford-Norwalk metropolitan area, about 50% of the students attend school districts where the average test score is higher than about 0.4, and about 50% attend districts where the average test score is lower than 0.4. The test score units are given in standard deviations, hence, the difference between the top and bottom metro areas exceeds one standard deviation. Although only 20 metropolitan areas are shown here, a geographic pattern immediately emerges: The best school districts are located in the Northeast and the Midwest, while the school districts with the lowest average test scores are located in the South and West. However, differences within metro areas are also large. For instance, in the Bridgeport-Stamford-Norwalk metropolitan area, the population-weighted IQR is about 80% of a standard deviation.

Focusing further on within metro area differences in test scores, I use the population-weighted IQR as a measure of the spread of the distribution of test scores within a metropolitan area. Figure 6.4 plots this measure of spread on the x-axis, with multigroup segregation shown on the y-axis. The figure shows a strong dependence between the spread of the test scores

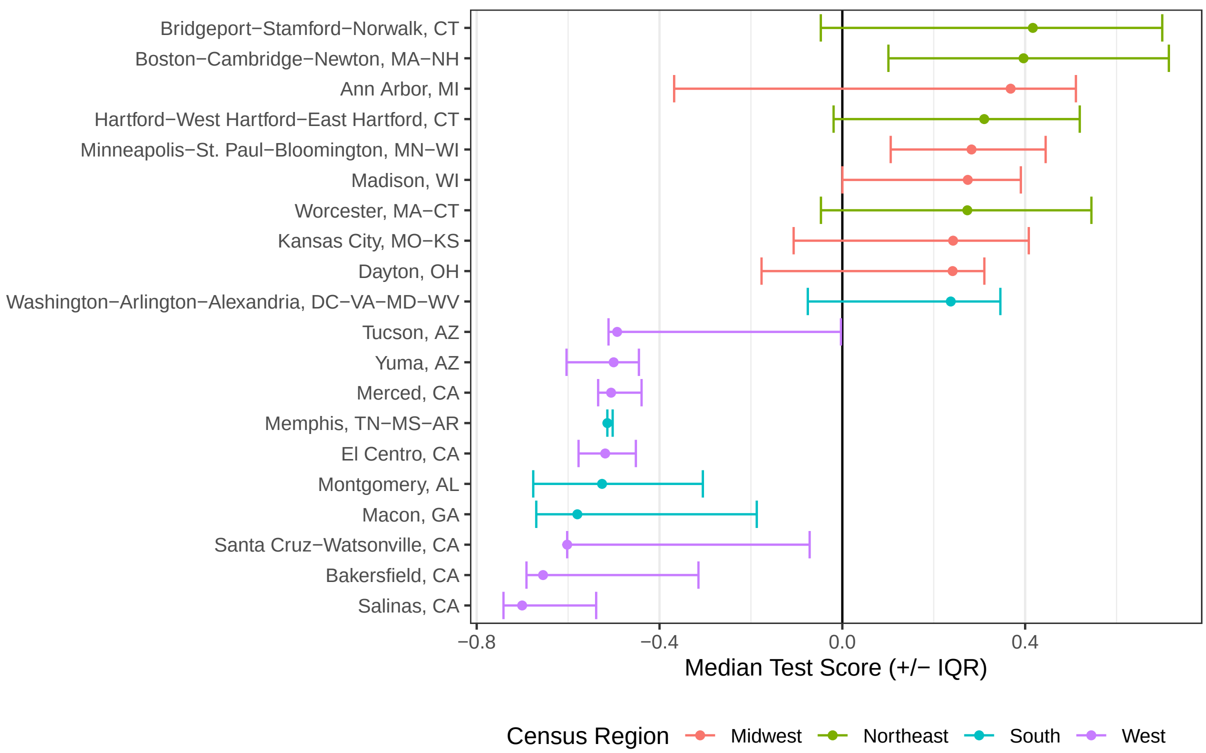

Figure 6.3: Top 10 and bottom 10 metro areas by test scores, 2010

Note: Based on calculations for 188 metropolitan areas. The quantiles shown are weighted by school district enrollment.

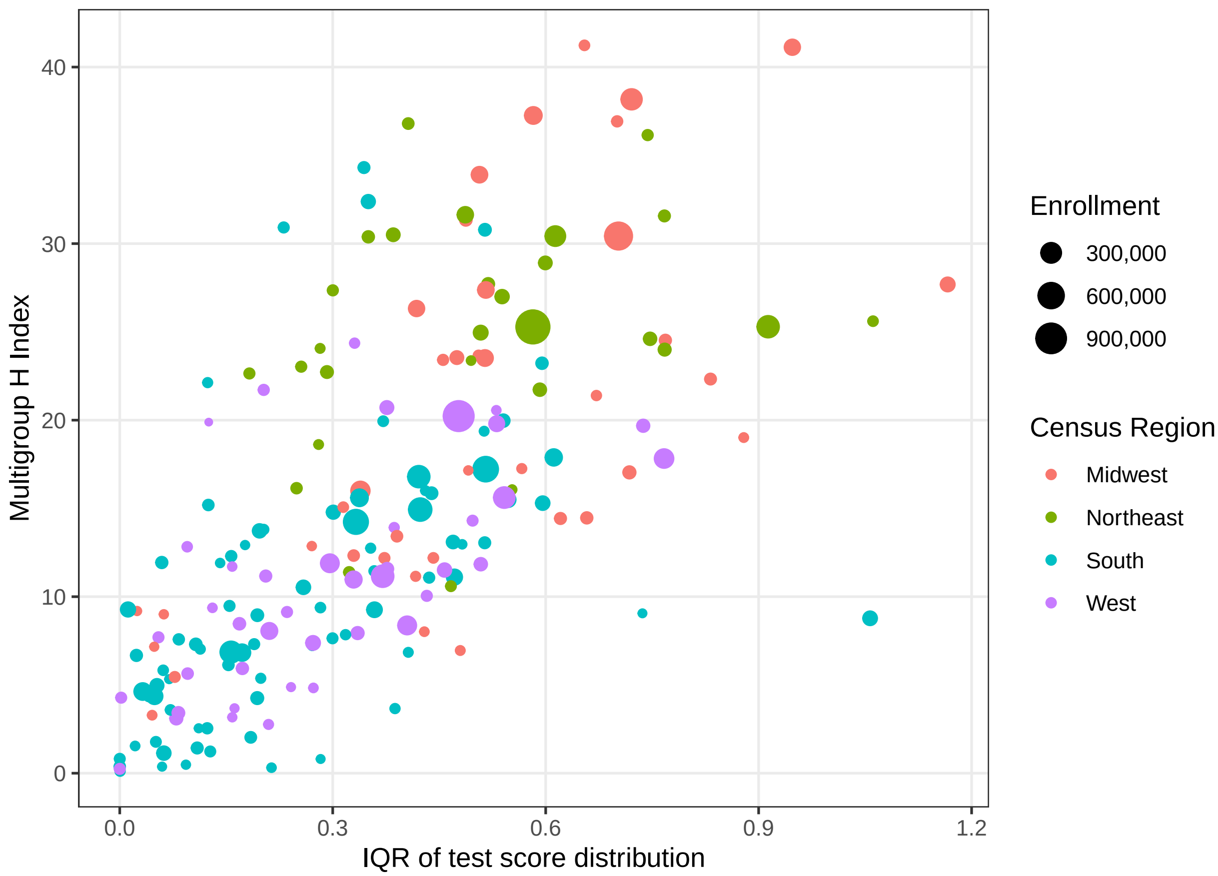

Figure 6.4: Test score spread and segregation

and segregation. (The correlation is 0.66). Some of the dependence is surely mechanical, as a higher number of districts will automatically increase both spread and segregation due to measurement and sampling error. Nonetheless, the strength of the relationship is suggestive of an important function of between-district school segregation, namely, to provide opportunities for the sorting of students by school district performance.

6.5.4 Cohort differences

For the remainder of the paper, I will switch to a cohort-centric perspective on segregation. The first question is how well a single cohort in a given year represents overall segregation in a metropolitan area. Overall, cohort differences are not expected to matter much (Fiel and Zhang 2018). If they largely represent (lagged) trends in fertility changes, one would expect only minor differences between, say, the racial composition of grades 3 and 8 for a given year. This is confirmed by computing correlations between cohort-specific and overall segregation. For instance, I take only the 2002 birth cohort in 2009 (when this cohort was in grade 4), compute the segregation based on this subset for the metropolitan area, and compare this value to the segregation of the metropolitan area in 2009, computed on the entire student population. Doing so for the three fully-observed birth cohorts (2001, 2002, and 2003), and for the years 2009-2016 yields correlations for each segregation index. For Black-White and Hispanic-White segregation, the correlations are all above 0.99, indicating that using only one cohort to represent overall segregation of the metropolitan area is not biasing the results. For Asian-White segregation, the correlations are also large, ranging from 0.97 to 0.98.

A subset of the comparisons are shown in Table 6.3, which also includes the change in segregation. This table focuses on the 2002 birth cohort, which is observed from 2010 (grade 3) to 2015 (grade 8). (Note that in Figure 6.1, the whole period 2009-2016 is observed.) The table also includes the segregation indices that are computed based on the subset of unified school districts, as described above. Overall, the magnitude of changes is similar between the 2002 birth cohort and all cohorts observed in those years. The largest difference is observed for Hispanic-White segregation in the Northeast (-2.4 for the 2002 cohort vs. -3.5 for all cohorts), which may reflect the larger demographic changes for the Hispanic group (esp. immigration) than for the other racial groups. Generally, as would be expected, the within-cohort changes are smaller than the overall changes, given that the latter incorporate cohort replacement *and* migration, while the former are assumed to almost exclusively reflect migration. The table also indicates that dropping non-unified school districts does not influence the levels and trends strongly.

6.5.5 Decomposition: Simple

As in Table 6.3, the decompositions focus on the 2002 birth cohort, taking it as broadly characteristic of the overall trends. I now decompose the changes in segregation for this birth cohort for the period 2010-2015, i.e. as this birth cohort moved from grade 3 to grade 8. The results are

Table 6.3: A cohort-centric view of segregation

	Multigroup			Black-White			Hispanic-White			Asian-White		
Subset	2010	2015	Δ	2010	2015	Δ	2010	2015	Δ	2010	2015	Δ
Midwest												
2002 cohort	27.0	25.2	-1.8	41.6	40.3	-1.4	23.6	21.0	-2.6	12.9	14.6	1.7
2002 cohort, unified	27.0	25.2	-1.8	41.6	40.2	-1.4	23.6	21.0	-2.6	12.9	14.6	1.7
All cohorts	26.9	24.7	-2.2	41.3	39.8	-1.5	22.8	20.6	-2.1	12.4	14.5	2.1
Northeast												
2002 cohort	26.3	24.6	-1.6	41.0	39.9	-1.1	34.6	32.2	-2.4	16.9	17.7	0.8
2002 cohort, unified	26.3	24.6	-1.6	40.7	39.9	-0.8	34.4	32.2	-2.2	16.6	17.5	1.0
All cohorts	26.1	24.0	-2.0	40.7	39.4	-1.3	34.6	31.0	-3.5	16.6	17.3	0.6
South												
2002 cohort	12.1	11.9	-0.2	18.5	19.0	0.6	12.8	12.4	-0.3	8.7	9.9	1.2
2002 cohort, unified	12.1	11.9	-0.2	18.5	19.0	0.6	12.8	12.4	-0.4	8.7	9.8	1.1
All cohorts	12.2	11.5	-0.8	18.4	18.4	0.0	12.4	12.1	-0.3	8.5	9.5	0.9
West												
2002 cohort	13.7	13.7	0.0	19.1	19.4	0.3	18.7	18.0	-0.7	11.7	12.2	0.5
2002 cohort, unified	13.7	13.7	0.0	19.1	19.4	0.4	18.5	17.9	-0.6	11.8	12.3	0.5
All cohorts	13.6	13.6	-0.1	19.3	18.8	-0.5	18.7	17.6	-1.1	11.7	11.8	0.1

then interpreted as changes in segregation that are brought about by residential mobility to and from school districts, i.e. the net change in racial group counts by school districts. This mobility includes international migration.

Figure 6.5 shows the results of this decomposition, separately for Multigroup, Black-White, Hispanic-White, and Asian-White segregation, and differentiated by Census region. The overall pattern markedly resembles the changes in residential segregation over the long term that were observed in the previous chapter. The Multigroup segregation decomposition captures the overall pattern: Residential mobility of the Black and Hispanic student populations lead to decreases in multigroup segregation, while residential mobility of the White population lead to increases in multigroup segregation. Residential mobility of the Asian population, likely owing to its small size, contributes very little to segregation change.

The dichotomous measures reveal some heterogeneity. For Black-White segregation, the small increases observed in the South and West, as well as the decreases observed in the Midwest and Northeast, are all the product of two large and countervailing forces: White residential mobility across school districts, on average, increased school segregation, while Black residential mobility decreased segregation. A very similar pattern is observed for Hispanic-White segregation, although the segregation-decreasing role of Hispanic residential mobility has a much larger effect. This is especially true in the Midwest, where Hispanic residential mobility alone would have decreased segregation by 5 points, which is a very large decline for a six-year period. Lastly, the size of the components is much smaller than for Asian-White segregation. This case also differs from the other two, as both the Asian and the White population contribute towards increases in segregation, except in the West.

While the descriptive results from the previous sections suggest some differences between U.S. regions in terms of segregation patterns, what is significant about Figure 6.5 is that the patterns of change are all similar, although they differ in magnitude: All contributions of the White population are positive, while the contributions of the Black and Hispanic population are always negative. These results again show that the absence of segregation change (e.g., Black-White segregation in the South) does not imply stability of the populations. A small change in segregation can hide large underlying changes in the distribution of the racial groups.

6.5.6 Decomposition: By racial composition

The first detailed decomposition is based on the racial group composition of the school district in 2010. Similarly to the last chapter, school districts are classified as "majority White" when the school district is more than 50% White. For the other groups, the criterion is identical. School districts are marked as "dominant White" when the school district is more than 90% White. All other district are "mixed." For the pairwise decomposition, districts that do not match the two racial groups for which the index is computed are also treated as mixed. For instance, a "majority Hispanic" district is treated as mixed when computing Black-White segregation. This reduces the number of factors and makes the results more interpretable. I also do not show the decomposition results for the multigroup index, due to the large number of factors (4 racial groups × 2 trajectories (growth/decline) × 6 district types = 48 factors).

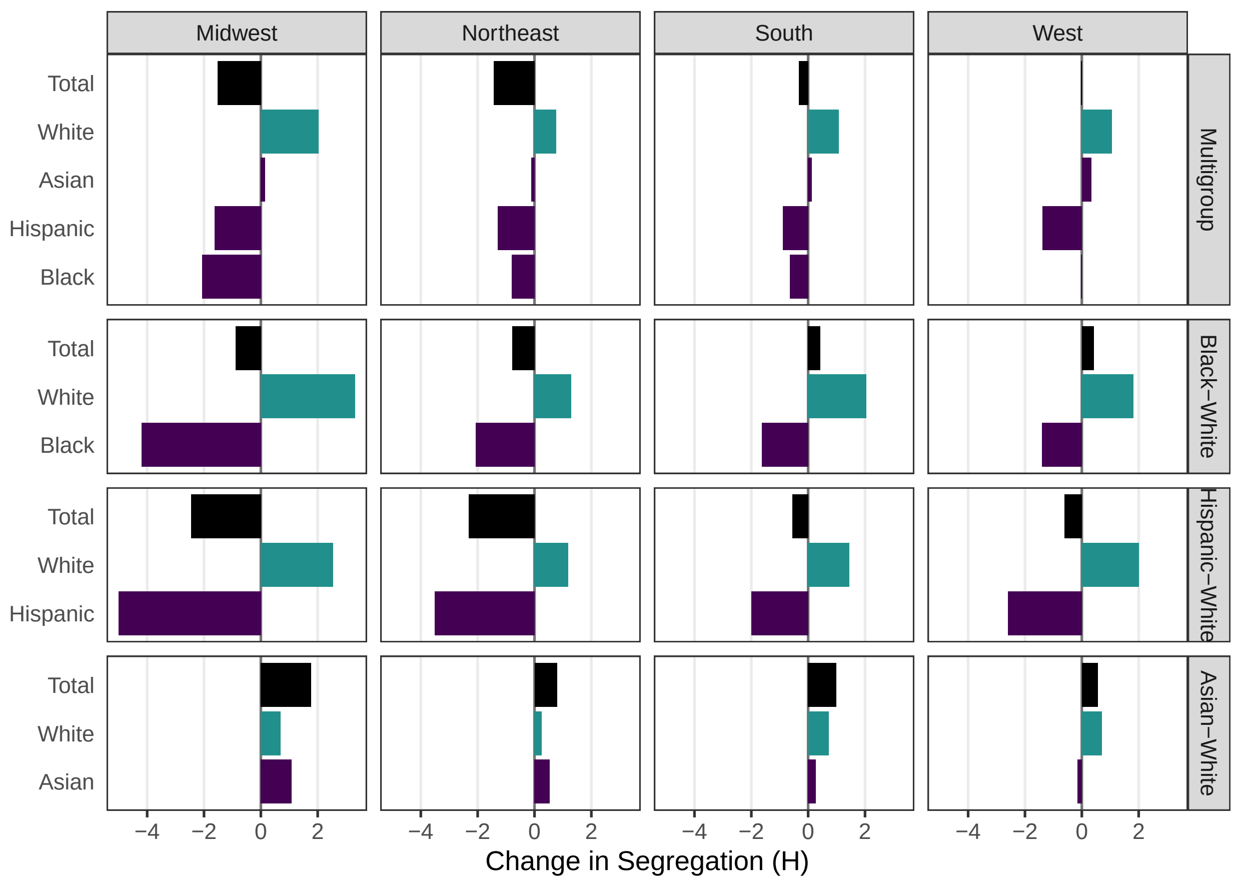

Figure 6.5: Simple decomposition by racial group, 2010-2015, for 2002 birth cohort

Before turning to the decomposition results, Table 6.4 shows the percentage changes for the different racial groups by school district composition in 2010, for the 2002 birth cohort only. Note that these numbers are all within-cohort changes, i.e. mostly attributed to residential mobility after grade 3. Over the course of five years, the White population increased modestly in dominant White school district, but declined strongly in majority Black, Hispanic, and mixed school districts. We would expect that these mobility flows increase segregation. The Black population generally declined in majority Black and Hispanic school districts, but increased in dominant White and majority White districts of the Midwest and Northeast. In the South and West, however, Black students mostly moved out of White school districts. The highest growth rates are observed for the Hispanic population. These are likely also a product of ongoing immigration. The number of Hispanic students increased almost everywhere, but especially strongly in dominant and majority White school districts. These movements likely led to strong declines in school segregation. The table also shows some migration of Asian families to White school districts, but here, the patterns are more differentiated by region. It is noticeable that the number of Asian students increased in all school district types in the Northeast. Again, immigration is likely a contributing factor to these developments.

Figure 6.6 shows the results for Black-White, Hispanic-White, and Asian-White segregation. Due to the increased complexity, results are not broken down by Census regions in this figure. Table 6.6 in the Appendix contains results that are broken down by region. For Black-White and Hispanic-White segregation, Figure 6.6 again shows that the residential mobility of the White population increases between-district segregation, which could be interpreted as ongoing white flight. The effect is strongest for mixed school districts, i.e. districts that are neither majority White nor Black, where the White population is moving out. However, White families moving out of majority Black and majority Hispanic school districts also contribute to increases in segregation. Another relatively large contributing factor that increases segregation is the mobility of Black families out of majority and dominant White school districts.

Table 6.4: Growth and decline (% change) of racial groups by school district composition, 2010-2015, 2002 birth cohort

Composition	White	Black	Hispanic	Asian
Midwest				
dom. White	3.7	20.7	63.5	6.7
maj. White	-0.9	12.2	29.5	6.0
maj. Black	-27.1	-15.0	-3.3	13.1
maj. Hispanic	-12.4	-13.9	-2.3	-
maj. Asian	-	-	-	-
Mixed	-17.0	-13.9	0.3	-0.8
Northeast				
dom. White	1.1	16.7	49.1	14.4
maj. White	-1.4	6.6	30.1	9.0
maj. Black	-14.6	-12.0	1.8	15.1
maj. Hispanic	-11.9	-12.4	2.0	4.4
maj. Asian	-4.8	-	-	4.3
Mixed	-6.1	4.0	11.2	8.0
South				
dom. White	1.3	-6.9	62.8	-
maj. White	2.1	1.8	29.1	13.2
maj. Black	-17.4	-8.7	16.9	-4.0
maj. Hispanic	-9.6	-13.6	2.6	-7.1
maj. Asian	-	-	-	-
Mixed	-6.6	2.0	14.5	8.7
West				
dom. White	3.2	-	-	-
maj. White	1.7	-7.9	21.3	2.3
maj. Black	-	-	-	-
maj. Hispanic	-9.5	-9.2	-2.5	0.6
maj. Asian	5.5	-	18.0	7.0
Mixed	-7.0	-5.8	8.7	2.2

Note: Percentage changes are not shown when the cell contained less than 500 students in 2010.

Regarding factors that decrease segregation, I find again that it is only the "minority" populations—Black and Hispanic—whose residential mobility decreases segregation. For Black-White segregation, the strongest factors are Black mobility to White school districts, but also outmigration from mixed and majority Black districts. For Hispanic-White segregation, the total decline is almost entirely driven by Hispanic mobility towards White school districts. This finding aligns well with the numbers in Table 6.4. For Asian-White segregation, the only relevant declines are again produced by Asian mobility to White districts, however, the segregation-increasing factors are larger here. While the decline of White students in mixed districts is the single factor that contributed the most to increases in segregation, the mobility of Asian families out of White districts and towards mixed districts all led to increases in segregation. The fact that there is significant mobility of Asian families both *in* and *out* of dominant and majority White school districts reveals that there exists significant heterogeneity in local circumstances both within and across metro areas that is only hinted at in this aggregate analysis.

While some individual factors are very large (exceeding 1 on the H scale), the total combined effect of all population movements is a net change of about zero for Black-White segregation. For Hispanic-White segregation, the residential mobility of Hispanics into majority White and dominant White districts is so large that the overall net effect of segregation change is large and negative. If we combine these two effects, the effect of Hispanic mobility to White school districts decreased segregation by about 4.5 points in the Midwest and Northeast, and by about 2 points in the South and West. For Asian-White segregation, there is a net increase in segregation. Here, the offsetting factors are small.

6.5.7 Decomposition: By test score distribution

In this section, I present another detailed decomposition that is based on the test score performance of school districts. To do so, I rank school districts by average 2010 test scores (across all students) within metropolitan area. I then split the districts into three groups, which are

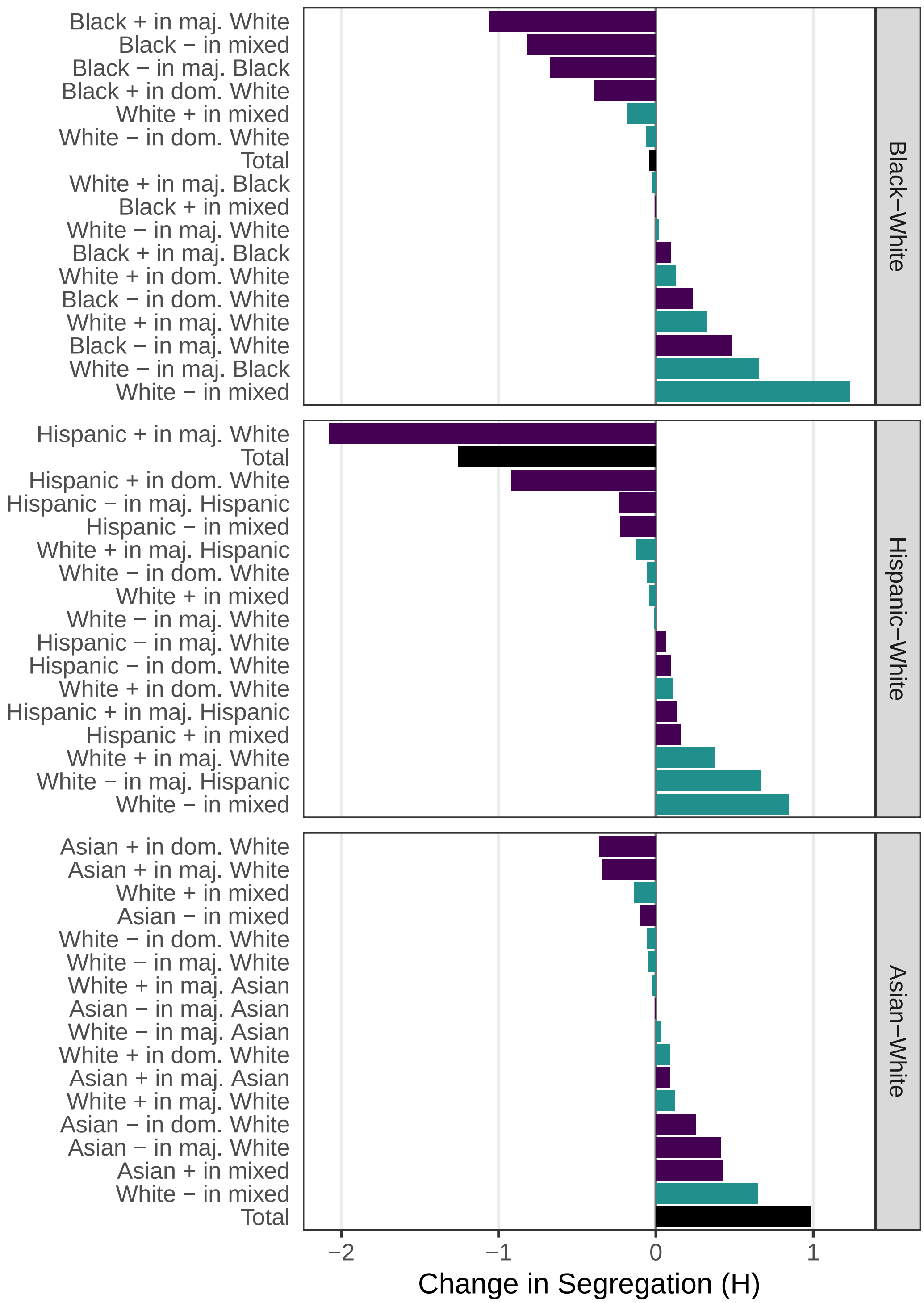

Figure 6.6: Decomposition by racial composition of school districts, 2010-2015, for 2002 birth cohort

weighted by the total enrollment of the 2002 birth cohort in the district. The resulting groups are formed such that roughly one third of the students are categorized to be in the top, middle, and bottom performing districts according to the test score distribution. This methodology excludes metropolitan areas that have fewer than three school districts. I also exclude metropolitan areas where it is not possible to reliably form three categories. This is the case when the distribution of students over school districts is highly skewed, for instance, when there are ten districts, but only two contain a substantial number of students. The decompositions that are presented in this section are therefore based on 151 school districts, compared to 193 in the preceding sections.

As would be expected, racial composition and test score performance are not independent. For 2010, the top performing districts are on average, 64% White, 19% Hispanic, 9% Black, and 8% Asian. For the districts in the bottom third these percentages change to 29% White, 34% Hispanic, 33% Black, and 4% Asian. It is therefore possible that the "racialized" patterns observed in the previous section are due to the close relationship between test score performance and school district racial composition.

Table 6.5 shows how the racial groups grew or declined between 2010 and 2015, conditional on test score performance of the school district. For White students, a clear pattern emerges: White families everywhere moved out of districts in the bottom third of the performance distribution, and moved into districts in the top third of the performance distribution. Similar patterns are observed for the Black population, where the growth rates in well-performing districts are very high in all regions, except the West. The Hispanic student population grows almost everywhere, but especially strongly in well-performing districts, and especially in the Midwest and Northeast. The pattern is more location-dependent for Asian students, likely also owing to somewhat higher variability due to lower sample sizes. Of course, these numbers do not represent causal effects. We do not know whether parents move into school districts because of their test score performance, and this relationship can well be confounded by a number of other variables (such as housing prices or job availability).

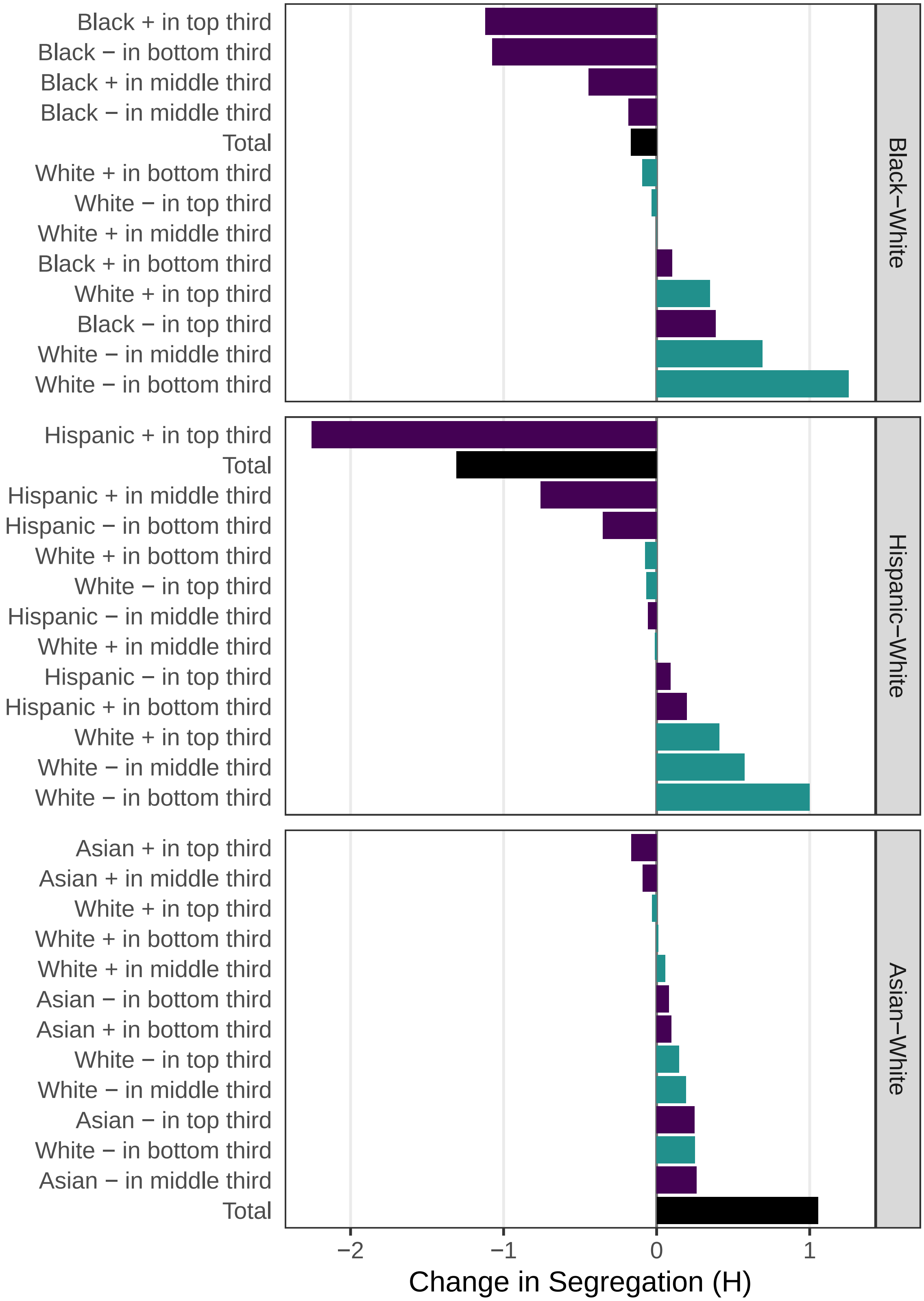

Figure 6.7: Decomposition by test score performance of school districts, 2010-2015, for 2002 birth cohort

Table 6.5: Growth and decline (% change) of racial groups by test score performance, 2010-2015, 2002 birth cohort

Test Scores	White	Black	Hispanic	Asian
Midwest				
bottom third	-14.5	-14.3	-0.1	6.2
middle third	-3.0	-6.2	7.7	-3.1
top third	3.6	18.5	43.6	8.9
Northeast				
bottom third	-8.3	-10.0	3.9	8.1
middle third	-2.6	5.6	14.5	7.0
top third	1.3	13.9	35.9	11.0
South				
bottom third	-9.9	-8.1	3.4	-2.6
middle third	-4.1	-4.7	8.9	2.1
top third	2.5	7.3	18.9	13.9
West				
bottom third	-7.7	-7.9	-1.7	-7.2
middle third	-6.0	-10.7	-0.2	-7.1
top third	2.8	-1.2	10.7	12.2

Given the patterns in Table 6.5, we would expect that the mobility patterns for the minority populations decrease segregation, and the mobility patterns for the White population increase segregation. The decomposition results are shown in Figure 6.7, and the results broken down by region are shown in Table 6.7 in the Appendix. For Black-White segregation, Figure 6.7 shows that the stratification by test score performance is an important predictor of residential mobility. Black mobility towards school districts in the top and middle third of the performance distribution, and from school districts in the bottom third of the distribution is associated with large declines in segregation. Similar movements of the White population, however, increased segregation. Of course, in total these movements have to balance each other out, as there has been only a small decline in Black-White segregation.

For Hispanic-White segregation, the pattern is very similar: Mobility of Hispanic students towards school districts in the top and middle third of the performance distribution is associated with large declines in segregation. White mobility from districts in the middle and bottom thirds of the performance distribution, and towards well-performing districts increased segregation. A major difference between Hispanic-White and Black-White segregation change with regards to test scores is that Hispanic mobility rates towards well-performing school districts have been very large, as Table 6.5 has shown. This effect alone would have led to decreases in segregation by more than two points on the H scale, a very large effect.

Unlike for Black-White and Hispanic-White segregation, where a small number of mobility flows are associated with large changes in segregation, the increase in Asian-White segregation was produced by many small movements. This seems to indicate that Asian-White segregation is less structured by test score performance differences than is the case for other racial groups.

6.6 Discussion and conclusion

This paper examined in detail the connection between residential mobility flows and between-district school segregation. The descriptive results show that Black-White and Hispanic-White school segregation declined in recent years especially in the Midwest and Northeast, with Hispanic-White segregation seeing especially strong declines. In the South and West declines were smaller or segregation stayed stable, however, average levels of segregation are much lower here than in the Midwest and Southeast. Asian-White segregation, on the other hand, increased everywhere. However, it is still at much lower average levels than Black-White and Hispanic-White segregation. One reason for the lower levels of between-district segregation in the South and West is the fact that school district fragmentation is much lower in these regions than in the Northeast and Midwest. The smaller number of districts in metropolitan areas of the South and West also makes it less likely that strong differences in average test score performance between school districts occur.

The core contribution of the paper is a series of over-time decompositions that quantify the effect of specific mobility flows on increases and decreases in segregation. Most of these decompositions reveal that there are large, (partially) offsetting effects, which shows that metropolitan area segregation is not fixed, but constantly remade. This is most clearly shown by the case of

Black-White segregation: Select mobility flows would have led to large increases or decreases in segregation, but the aggregate effect is one of very small segregation change. Hispanic-White segregation also shows a similar partial offsetting, however, in this case the declines produced by Hispanic mobility and immigration were much larger than the segregation-reinforcing mobility of the White population. One could (again) summarize these patterns as simultaneous resegregation and desegregation: The minority populations, including to some extent the Asian population, move into majority White school districts, while White families move out of majority Hispanic, majority Black, and mixed districts. Although this chapter studied a much more recent time period (2010-2015) compared to the previous chapter (1990-2010), the fundamental relationships seem to be unchanged.

Finally, this paper also presented a decomposition that was not based on the racial composition of the school districts, but based on their test score performance. Given the strong dependency between test score performance and racial composition, it is not surprising that the roles of the racial groups in producing or diminishing segregation are similar. However, one can now also see the similarities between the groups: White, Black, and Hispanic families *all* move from school districts with low average test scores to school districts with high average test scores. If school district performance is really as important a driver of residential location decisions as suggested by this result, the "offsetting" character of residential segregation change in the U.S. can possibly be explained by the large differences in school district performance.

As I studied only aggregate data in this paper, I can only speculate on the motivations for individual families to move. However, I would argue that the results shown in this paper call for a deeper engagement of sociology with residential mobility, a topic that is much more prominent in neighboring disciplines such as demography, economics (for a review, see Brunner 2014), and (especially) geography (e.g., Clark and Onaka 1983; Galster 1988).[4] This is despite the fact that neighborhoods and neighborhood change are central topics in sociology.

[4]Notable recent exceptions in sociology are the works of Kyle Crowder and Maria Krysan (Crowder 2000; Crowder and South 2008; Krysan and Crowder 2017), and Elizabeth Bruch and Rob Mare (Bruch and Mare 2006, 2012).

An important extension of this work is therefore to use longitudinal microdata to better understand the residential mobility flows that were identified as important through the decompositions. One interesting aspects of a longitudinal dataset is that, unlike in this paper, both the *source* and the *destination* of the mobility flow could be observed. This would allow even more precise statements about the mobility flows that shape segregation. Another important contribution of a longitudinal dataset is that residential mobility before and after schooling is observed. In this paper, I only observe mobility while children are in school. Cadwallader (1992) argues that during this period, households are less mobile compared to the period before children enter school and after children leave the household (see also Clark and Onaka 1983). Hence, the numbers presented in this paper can be argued to be lower bounds. If one would include the mobility of the cohort before they started school, the patterns would likely be strengthened.

Lastly, detailed longitudinal microdata could be used to understand the fundamental causes of family residential mobility in the first place, completing the link between micro-level decision-making and macro-level outcomes, i.e. segregation. While schools play a large role in moving decisions, there clearly are other factors that both help and impede relocation decisions. Furthermore, not all residential mobility is voluntary—for instance, economically struggling families may have to move out of desirable school districts (Schafft 2006). Future work on understanding segregation therefore needs to more explicitly engage with the fundamental problem of using decisions and constraints at the micro level to explain macro level outcomes—both theoretically and methodologically.

6.7 Appendix

Table 6.6: School district racial composition results by region

Factor	Midwest	Northeast	South	West
Black-White				
Black - in dom. White	0.5	0.5	0.1	0.0
Black + in dom. White	-0.9	-1.0	-0.1	-0.1
Black - in maj. Black	-1.7	-1.0	-0.5	0.0
Black + in maj. Black	0.2	0.1	0.1	0.0
Black - in maj. White	0.3	0.7	0.4	0.6
Black + in maj. White	-1.6	-1.3	-0.9	-0.6
Black - in mixed	-1.0	-0.4	-0.5	-1.3
Black + in mixed	0.0	0.3	-0.1	0.0
White - in dom. White	-0.1	-0.1	0.0	0.0
White + in dom. White	0.4	0.2	0.1	0.0
White - in maj. Black	1.5	0.6	0.7	0.0
White + in maj. Black	0.0	0.0	0.0	0.0
White - in maj. White	0.3	0.0	0.0	-0.1
White + in maj. White	0.1	0.2	0.4	0.4
White - in mixed	1.2	1.0	1.0	1.8
White + in mixed	0.0	-0.5	0.0	-0.3
Hispanic-White				
Hispanic - in dom. White	0.2	0.3	0.0	0.0
Hispanic + in dom. White	-2.4	-1.8	-0.3	-0.2
Hispanic - in maj. Hispanic	-0.1	-0.1	-0.2	-0.5
Hispanic + in maj. Hispanic	0.0	0.2	0.1	0.1
Hispanic - in maj. White	0.0	0.2	0.0	0.0
Hispanic + in maj. White	-2.3	-2.7	-1.9	-1.7
Hispanic - in mixed	-0.9	-0.2	-0.1	0.0
Hispanic + in mixed	0.4	0.6	0.2	-0.4
White - in dom. White	-0.1	-0.1	0.0	0.0
White + in dom. White	0.3	0.1	0.1	0.0
White - in maj. Hispanic	0.2	0.5	0.5	1.4
White + in maj. Hispanic	0.0	-0.2	0.0	-0.3
White - in maj. White	0.3	0.0	-0.1	-0.1
White + in maj. White	0.1	0.2	0.4	0.7
White - in mixed	1.8	1.0	0.7	0.2
White + in mixed	0.0	-0.3	0.0	0.1
Asian-White				

(Continued on next page...)

Table 6.6: School district racial composition results by region *(continued)*

Factor	Midwest	Northeast	South	West
Asian - in dom. White	0.7	0.4	0.1	0.0
Asian + in dom. White	-0.8	-0.8	-0.1	-0.1
Asian - in maj. Asian	0.0	0.0	0.0	0.0
Asian + in maj. Asian	0.0	0.1	0.0	0.3
Asian - in maj. White	0.5	0.4	0.3	0.4
Asian + in maj. White	0.3	-0.4	-0.4	-0.7
Asian - in mixed	-0.1	-0.1	0.0	-0.2
Asian + in mixed	0.4	1.0	0.4	0.1
White - in dom. White	-0.1	-0.1	0.0	0.0
White + in dom. White	0.2	0.1	0.0	0.0
White - in maj. Asian	0.0	0.0	0.0	0.1
White + in maj. Asian	0.0	0.0	0.0	-0.1
White - in maj. White	0.1	0.0	-0.1	-0.1
White + in maj. White	-0.1	0.0	0.2	0.3
White - in mixed	0.6	0.5	0.6	0.8
White + in mixed	0.0	-0.2	0.0	-0.3

Table 6.7: Test score decomposition results by region

Factor	Midwest	Northeast	South	West
Black-White				
Black - in bottom third	-1.8	-1.2	-0.8	-0.7
Black + in bottom third	0.1	0.2	0.0	0.1
Black - in middle third	-0.5	0.5	-0.2	-0.5
Black + in middle third	-1.0	-0.7	-0.3	0.0
Black - in top third	0.4	0.6	0.4	0.3
Black + in top third	-1.4	-1.4	-1.0	-0.9
White - in bottom third	2.0	1.3	1.0	0.9
White + in bottom third	0.0	-0.4	0.0	0.0
White - in middle third	1.0	0.3	0.7	0.8
White + in middle third	0.1	-0.1	0.1	-0.2
White - in top third	-0.1	-0.1	0.0	0.0
White + in top third	0.3	0.3	0.4	0.4
Hispanic-White				
Hispanic - in bottom third	-0.8	-0.2	-0.2	-0.3
Hispanic + in bottom third	0.2	0.3	0.2	0.1
Hispanic - in middle third	-0.1	0.1	-0.1	-0.2
Hispanic + in middle third	-1.4	-1.1	-0.4	-0.4

(Continued on next page...)

Table 6.7: Test score decomposition results by region *(continued)*

Factor	Midwest	Northeast	South	West
Hispanic - in top third	0.1	0.2	0.0	0.1
Hispanic + in top third	-3.2	-2.7	-1.7	-1.8
White - in bottom third	1.4	1.2	0.8	0.8
White + in bottom third	0.0	-0.3	0.0	0.0
White - in middle third	0.9	0.4	0.4	0.7
White + in middle third	0.1	-0.1	0.1	-0.2
White - in top third	-0.1	-0.1	-0.1	0.0
White + in top third	0.3	0.2	0.4	0.7
Asian-White				
Asian - in bottom third	0.2	0.0	0.1	0.0
Asian + in bottom third	0.2	0.4	0.0	-0.1
Asian - in middle third	0.6	0.4	0.1	0.1
Asian + in middle third	-0.3	0.1	-0.1	0.0
Asian - in top third	0.4	0.3	0.2	0.1
Asian + in top third	0.1	-0.7	0.1	-0.3
White - in bottom third	0.3	0.4	0.2	0.2
White + in bottom third	0.0	-0.1	0.0	0.1
White - in middle third	0.2	0.0	0.2	0.3
White + in middle third	0.1	0.0	0.1	0.0
White - in top third	0.1	0.0	0.1	0.3
White + in top third	-0.1	0.0	0.1	-0.2

References

Agresti, Alan (2013). *Categorical Data Analysis*. Third Edition. Hoboken, NJ: Wiley.

Allmendinger, Jutta (1989). “Educational Systems and Labor Market Outcomes”. In: *European Sociological Review* 5.3, pp. 231–250. DOI: 10.1093/oxfordjournals.esr.a036524.

An, Brian P. and Adam Gamoran (2009). “Trends in School Racial Composition in the Era of Unitary Status”. In: *From the Courtroom to the Classroom*. Ed. by Claire Smrekar and Ellen Goldring. Cambridge, MA: Harvard Education Press, pp. 19–47.

Andrews, K. T. (2002). “Movement-Countermovement Dynamics and the Emergence of New Institutions: The Case of "White Flight" Schools in Mississippi”. In: *Social Forces* 80.3, pp. 911–936. DOI: 10.1353/sof.2002.0001.

Aumann, Robert J. (1994). “The Shapley Value”. In: *Game-Theoretic Methods in General Equilibrium Analysis*. Ed. by Jean-François Mertens and Sylvain Sorin. Dordrecht: Springer Netherlands, pp. 61–66. DOI: 10.1007/978-94-017-1656-7_6.

Bader, Michael and Siri Warkentien (2016). “The Fragmented Evolution of Racial Integration since the Civil Rights Movement”. In: *Sociological Science* 3, pp. 135–164. DOI: 10.15195/v3.a8.

Benavot, Aaron (1983). “The Rise and Decline of Vocational Education”. In: *Sociology of Education* 56.2, pp. 63–76. DOI: 10.2307/2112655.

Bhrolcháin, Máire Ní and Tim Dyson (2007). “On Causation in Demography: Issues and Illustrations”. In: *Population and Development Review* 33.1, pp. 1–36. DOI: 10.1111/j.1728-4457.2007.00157.x.

Bifulco, Robert and Helen F. Ladd (2007). “School Choice, Racial Segregation, and Test-Score Gaps: Evidence from North Carolina’s Charter School Program”. In: *Journal of Policy Analysis and Management* 26.1, pp. 31–56. DOI: 10.1002/pam.20226.

Bischoff, Kendra (2008). “School District Fragmentation and Racial Residential Segregation: How Do Boundaries Matter?” In: *Urban Affairs Review* 44.2, pp. 182–217. DOI: 10.1177/1078087408320651.

Bol, Thijs, Christina Ciocca Eller, Herman G. van de Werfhorst, and Thomas A. DiPrete (2019). “School-to-Work Linkages, Educational Mismatches, and Labor Market Outcomes”. In: *American Sociological Review* 84.2, pp. 275–307. DOI: 10.1177/0003122419836081.

Bol, Thijs and Herman G. Van de Werfhorst (2013). "Educational Systems and the Trade-off between Labor Market Allocation and Equality of Educational Opportunity". In: *Comparative Education Review* 57.2, pp. 285–308. DOI: 10.1086/669122.

Brauns, Hildegard, Susanne Steinmann, Annick Kieffer, and Catherine Marry (1999). "Does Education Matter? France and Germany in Comparative Perspective". In: *European Sociological Review* 15.1, pp. 61–89. DOI: 10.1093/oxfordjournals.esr.a018255.

Breen, Richard (2005). "Explaining Cross-National Variation in Youth Unemployment". In: *European Sociological Review* 21.2, pp. 125–134. DOI: 10.1093/esr/jci008.

Bruch, Elizabeth E. and Robert D. Mare (2006). "Neighborhood Choice and Neighborhood Change". In: *American Journal of Sociology* 112.3, pp. 667–709. DOI: 10.1086/507856.

— (2012). "Methodological Issues in the Analysis of Residential Preferences, Residential Mobility, and Neighborhood Change". In: *Sociological Methodology* 42.1, pp. 103–154. DOI: 10.1177/0081175012444105.

Brunner, Eric J. (2014). "School Quality, School Choice, and Residential Mobility". In: *Education, Land, and Location*. Ed. by Gregory K. Ingram and Daphne A. Kenyon. Cambridge, MA: Lincoln Institute of Land Policy, pp. 62–87.

Bundesinstitut für Berufsbildung [BIBB] (2006). "BIBB/BAuA-Erwerbstätigenbefragung 2006". In: DOI: 10.7803/501.06.1.1.30.

Cadwallader, Martin T. (1992). *Migration and Residential Mobility: Macro and Micro Approaches*. Madison: University of Wisconsin Press.

Card, David, Jörg Heining, and Patrick Kline (2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality". In: *The Quarterly Journal of Economics* 128.3, pp. 967–1015. DOI: 10.1093/qje/qjt006.

Carrington, William J. and Kenneth R. Troske (1998). "Interfirm Segregation and the Black-/White Wage Gap". In: *Journal of Labor Economics* 16.2, pp. 231–260. DOI: 10.1086/209888.

Charles, Camille Zubrinsky (2003). "The Dynamics of Racial Residential Segregation". In: *Annual Review of Sociology* 29.1, pp. 167–207. DOI: 10.1146/annurev.soc.29.010202.100002.

Charles, Maria (1992). "Cross-National Variation in Occupational Sex Segregation". In: *American Sociological Review* 57.4, pp. 483–502. DOI: 10.2307/2096096.

Charles, Maria and David B. Grusky (1995). "Models for Describing the Underlying Structure of Sex Segregation". In: *American Journal of Sociology* 100.4, pp. 931–971. DOI: 10.1086/230605.

Charles, Maria and David B. Grusky (2004). *Occupational Ghettos: The Worldwide Segregation of Women and Men*. Stanford, CA: Stanford University Press.

Clark, W. A. V. (1986). "Residential Segregation in American Cities: A Review and Interpretation". In: *Population Research and Policy Review* 5.2, pp. 95–127. DOI: 10.1007/BF00137176.

Clark, W. A. V. and Jun L. Onaka (1983). "Life Cycle and Housing Adjustment as Explanations of Residential Mobility". In: *Urban Studies* 20.1, pp. 47–57. DOI: 10.1080/713703176.

Clotfelter, Charles T. (2001). "Are Whites Still Fleeing? Racial Patterns and Enrollment Shifts in Urban Public Schools, 1987-1996". In: *Journal of Policy Analysis and Management* 20.2, pp. 199–221. DOI: 10.1002/pam.2022.

— (2011). *After "Brown": The Rise and Retreat of School Desegregation*. Princeton, NJ: Princeton University Press.

Coleman, James S., Ernest Q. Campbell, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York (1966). *Equality of Educational Opportunity*. Washington D.C.: U.S. Department of Health, Education, and Welfare.

Coleman, James S., Thomas Hoffer, and Sally Kilgore (1982a). "Achievement and Segregation in Secondary Schools: A Further Look at Public and Private School Differences". In: *Sociology of Education* 55.2/3, pp. 162–182. DOI: 10.2307/2112296.

— (1982b). *High School Achievement: Public, Catholic, and Private Schools Compared*. New York, NY: Basic Books.

Coleman, James S., Sara D. Kelly, and John A. Moore (1975). *Trends in School Segregation, 1968-73*. Washington, D.C: The Urban Institute.

Crowder, Kyle (2000). "The Racial Context of White Mobility: An Individual-Level Assessment of the White Flight Hypothesis". In: *Social Science Research* 29.2, pp. 223–257. DOI: 10.1006/ssre.1999.0668.

Crowder, Kyle and Scott J. South (2008). "Spatial Dynamics of White Flight: The Effects of Local and Extralocal Racial Conditions on Neighborhood Out-Migration". In: *American Sociological Review* 73.5, pp. 792–812. DOI: 10.1177/000312240807300505.

Day, Charles R. (2001). *Schools and Work. Technical and Vocational Education in France since the Third Republic*. Montreal & Kingston: McGill-Queen's University Press.

Deming, W. Edwards and Frederick F. Stephan (1940). "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known". In: *Annals of Mathematical Statistics* 11.4, pp. 427–444. DOI: 10.1214/aoms/1177731829.

Deutsch, Joseph, Yves Flückiger, and Jacques Silber (2009). “Analyzing Changes in Occupational Segregation: The Case of Switzerland (1970–2000)”. In: *Occupational and Residential Segregation*. Vol. 17. Research on Economic Inequality. Bingley, UK: Emerald, pp. 171–202.

DiPrete, Thomas A., Thijs Bol, Christina Ciocca Eller, and Herman G. van de Werfhorst (2017). “School-to-Work Linkages in the United States, Germany, and France”. In: *American Journal of Sociology* 122.6, pp. 1869–1938. DOI: 10.1086/691327.

Eurostat (n.d.). *European Labor Force Survey, 2005-2010*.

Farley, Reynolds and William H. Frey (1994). “Changes in the Segregation of Whites from Blacks During the 1980s: Small Steps Toward a More Integrated Society”. In: *American Sociological Review* 59.1, pp. 23–45. DOI: 10.2307/2096131.

Farley, Reynolds, Toni Richards, and Clarence Wurdock (1980). “School Desegregation and White Flight: An Investigation of Competing Models and Their Discrepant Findings”. In: *Sociology of Education* 53.3, pp. 123–139. DOI: 10.2307/2112408.

Farrell, Chad R. (2008). “Bifurcation, Fragmentation or Integration? The Racial and Geographical Structure of US Metropolitan Segregation, 1990-2000”. In: *Urban Studies* 45.3, pp. 467–499. DOI: 10.1177/0042098007087332.

Ferguson, John-Paul and Rembrand Koning (2018). “Firm Turnover and the Return of Racial Establishment Segregation”. In: *American Sociological Review* 83.3, pp. 445–474. DOI: 10.1177/0003122418767438.

Fiel, Jeremy E. (2013). “Decomposing School Resegregation: Social Closure, Racial Imbalance, and Racial Isolation”. In: *American Sociological Review* 78.5, pp. 828–848. DOI: 10.1177/0003122413496252.

Fiel, Jeremy E. and Yongjun Zhang (2018). “Three Dimensions of Change in School Segregation: A Grade-Period-Cohort Analysis”. In: *Demography* 55.1, pp. 33–58. DOI: 10.1007/s13524-017-0632-9.

Fischer, Claude S., Gretchen Stockmayer, Jon Stiles, and Michael Hout (2004). “Distinguishing the Geographic Levels and Social Dimensions of U.S. Metropolitan Segregation, 1960-2000”. In: *Demography* 41.1, pp. 37–59. DOI: 10.1353/dem.2004.0002.

Fischer, Mary J. (2008). “Shifting Geographies: Examining the Role of Suburbanization in Blacks’ Declining Segregation”. In: *Urban Affairs Review* 43.4, pp. 475–496. DOI: 10.1177/1078087407305499.

Forster, Andrea G. and Thijs Bol (2018). “Vocational Education and Employment over the Life Course Using a New Measure of Occupational Specificity”. In: *Social Science Research* 70, pp. 176–197. DOI: 10.1016/j.ssresearch.2017.11.004.

Forster, Andrea G., Thijs Bol, and Herman G. van de Werfhorst (2016). "Vocational Education and Employment over the Life Cycle". In: *Sociological Science* 3, pp. 473–494. DOI: 10.15195/v3.a21.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo (2011). "Decomposition Methods in Economics". In: *Handbook of Labor Economics*. Ed. by Orley Ashenfelter and David Card. Vol. 4. Elsevier, pp. 1–102. DOI: 10.1016/S0169-7218(11)00407-2.

Fossett, Mark (2017). *New Methods for Measuring and Analyzing Segregation*. Vol. 42. The Springer Series on Demographic Methods and Population Analysis. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-41304-4.

Fowler, Christopher S., Barrett A. Lee, and Stephen A. Matthews (2016). "The Contributions of Places to Metropolitan Ethnoracial Diversity and Segregation: Decomposing Change Across Space and Time". In: *Demography* 53.6, pp. 1955–1977. DOI: 10.1007/s13524-016-0517-3.

Frankel, David M. and Oscar Volij (2011). "Measuring School Segregation". In: *Journal of Economic Theory* 146.1, pp. 1–38. DOI: 10.1016/j.jet.2010.10.008.

Frankenberg, Erica, Chungmei Lee, and Gary Orfield (2003). *A Multiracial Society with Segregated Schools: Are We Losing the Dream?* Cambridge, MA: The Civil Rights Project, Harvard University.

Frankenberg, Erica, Genevieve Siegel-Hawley, and Jia Wang (2010). *Choice without Equity: Charter School Segregation and the Need for Civil Rights Standards*. Los Angeles, CA: The Civil Rights Project.

Frey, William H. (1979). "Central City White Flight: Racial and Nonracial Causes". In: *American Sociological Review* 44.3, pp. 425–448. DOI: 10.2307/2094885.

— (2018). *Diversity Explosion: How New Racial Demographics Are Remaking America*. Washington, D.C: Brookings Institution Press.

Galster, George (1988). "Residential Segregation in American Cities: A Contrary Review". In: *Population Research and Policy Review* 7.2, pp. 93–112. DOI: 10.1007/BF00125462.

Galster, George and Patrick Sharkey (2017). "Spatial Foundations of Inequality: A Conceptual Model and Empirical Overview". In: *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3.2, pp. 1–33. DOI: 10.7758/rsf.2017.3.2.01.

Glaeser, Edward and Jacob Vigdor (2012). *The End of the Segregated Century: Racial Separation in America's Neighborhoods, 1890-2010*. 66. Manhattan Institute.

Goux, Dominique and Eric Maurin (1998). "From Education to First Job: The French Case". In: *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. Ed. by Yossi Shavit and Walter Müller. Oxford: Clarendon Press, pp. 369–406.

Grusky, David B. and Maria Charles (1998). "The Past, Present, and Future of Sex Segregation Methodology". In: *Demography* 35.4, pp. 497–504. DOI: 10.2307/3004017.

Hall, Matthew (2013). "Residential Integration on the New Frontier: Immigrant Segregation in Established and New Destinations". In: *Demography* 50.5, pp. 1873–1896. DOI: 10.1007/s13524-012-0177-x.

Hall, Peter A. and David W. Soskice, eds. (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.

Haller, Max, Wolfgang König, Peter Krause, and Karin Kurz (1985). "Patterns of Career Mobility and Structural Positions in Advanced Capitalist Societies: A Comparison of Men in Austria, France, and the United States". In: *American Sociological Review* 50.5, pp. 579–603. DOI: 10.2307/2095376.

Hanushek, Eric A., Guido Schwerdt, Ludger Woessmann, and Lei Zhang (2017). "General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle". In: *Journal of Human Resources* 52.1, pp. 48–87. DOI: 10.3368/jhr.52.1.0415-7074R.

Iceland, John (2004). "Beyond Black and White". In: *Social Science Research* 33.2, pp. 248–271. DOI: 10.1016/S0049-089X(03)00056-5.

INSEE/ADISP-CMH (n.d.[a]). *Enquêtes Emploi, Série 1990-2002*.

— (n.d.[b]). *Formation, Qualification Professionnelle (FQP), 1970 and 1985*.

James, David R. and Karl E. Taeuber (1985). "Measures of Segregation". In: *Sociological Methodology* 15, pp. 1–32. DOI: 10.2307/270845.

Jann, Ben (2008). "The Blinder–Oaxaca Decomposition for Linear Regression Models". In: *The Stata Journal* 8.4, pp. 453–479. DOI: 10.1177/1536867X0800800401.

Joseph, Andreas (2020). *Parametric Inference with Universal Function Approximators*. arXiv 1903.04209.

Karmel, T. and M. Maclachlan (1988). "Occupational Sex Segregation —Increasing or Decreasing?" In: *Economic Record* 64.3, pp. 187–195. DOI: 10.1111/j.1475-4932.1988.tb02057.x.

Kerckhoff, Alan C. (1995). "Institutional Arrangements and Stratification Processes in Industrial Societies". In: *Annual Review of Sociology* 21.1, pp. 323–347. DOI: 10.1146/annurev.so.21.080195.001543.

Kitagawa, Evelyn M. (1955). “Components of a Difference Between Two Rates”. In: *Journal of the American Statistical Association* 50.272, pp. 1168–1194. DOI: 10.2307/2281213.

König, Wolfgang and Walter Müller (1986). “Educational Systems and Labour Markets as Determinants of Worklife Mobility in France and West Germany: A Comparison of Men’s Career Mobility, 1965–1970”. In: *European Sociological Review* 2.2, pp. 73–96. DOI: 10.1093/oxfordjournals.esr.a036414.

Korver-Glenn, Elizabeth (2018). “Compounding Inequalities: How Racial Stereotypes and Discrimination Accumulate across the Stages of Housing Exchange”. In: *American Sociological Review* 83.4, pp. 627–656. DOI: 10.1177/0003122418781774.

Krysan, Maria and Kyle Crowder (2017). *Cycle of Segregation: Social Processes and Residential Stratification.* New York: Russell Sage Foundation.

Ladd, Helen, Charles T. Clotfelter, and John B. Holbein (2015). *The Growing Segmentation of the Charter School Sector in North Carolina.* Cambridge, MA: National Bureau of Economic Research.

Lee, Barrett A., Glenn Firebaugh, Stephen A. Matthews, Sean F. Reardon, Chad R. Farrell, and David O’Sullivan (2008). “Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographic Scales”. In: *American Sociological Review* 73.5, pp. 766–791. DOI: 10.1177/000312240807300504.

Lichter, Daniel T., Domenico Parisi, and Michael C. Taquino (2015). “Toward a New Macro-Segregation? Decomposing Segregation within and between Metropolitan Cities and Suburbs”. In: *American Sociological Review* 80.4, pp. 843–873. DOI: 10.1177/0003122415588558.

Loewen, James W. (2018). *Sundown Towns: A Hidden Dimension of American Racism.* New York: The New Press.

Logan, John R. (2004). *Resegregation in American Public Schools? Not in the 1990s.* Albany, NY: Lewis Mumford Center for Comparative Urban and Regional Research.

— (2016). “As Long As There Are Neighborhoods”. In: *City & Community* 15.1, pp. 23–28. DOI: 10.1111/cico.12149.

Logan, John R. and Deirdre Oakley (2004). *The Continuing Legacy of the Brown Decision.* Albany, NY: Lewis Mumford Center for Comparative Urban and Regional Research.

Logan, John R., Deirdre Oakley, and Jacob Stowell (2008). “School Segregation in Metropolitan Regions, 1970–2000: The Impacts of Policy Choices on Public Education”. In: *American Journal of Sociology* 113.6, pp. 1611–1644. DOI: 10.1086/587150.

Logan, John R., Brian J. Stults, and Reynolds Farley (2004). “Segregation of Minorities in the Metropolis: Two Decades of Change”. In: *Demography* 41.1, pp. 1–22. DOI: 10.1353/dem.2004.0007.

Logan, John R., Zengwang Xu, and Brian J. Stults (2014). “Interpolating U.S. Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database”. In: *The Professional Geographer* 66.3, pp. 412–420. DOI: 10.1080/00330124.2014.905156.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles (2019). “IPUMS National Historical Geographic Information System: Version 14.0 [Database]”. In: DOI: 10.18128/D050.V14.0.

Marry, Catherine, Annick Kieffer, Hildegard Brauns, and Susanne Steinmann (1998). “France-Allemagne: Inégales Avancées Des Femmes: Évolutions Comparées de l’éducation et de l’activité Des Femmes de 1971 à 1991”. In: *Revue française de sociologie* 39.2, pp. 353–389. DOI: 10.2307/3322720.

Massey, Douglas S. and Nancy A. Denton (1989). “Hypersegregation in U.S. Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions”. In: *Demography* 26.3, pp. 373–391. DOI: 10.2307/2061599.

Massey, Douglas S., Jonathan Rothwell, and Thurston Domina (2009). “The Changing Bases of Segregation in the United States”. In: *The ANNALS of the American Academy of Political and Social Science* 626.1, pp. 74–90. DOI: 10.1177/0002716209343558.

Maurice, Marc, François Sellier, and Jean-Jacques Silvestre (1986). *The social foundations of industrial power: a comparison of France and Germany*. Cambridge, MA: MIT Press.

Mora, Ricardo and Javier Ruiz-Castillo (2003). “Additively Decomposable Segregation Indexes. The Case of Gender Segregation by Occupations and Human Capital Levels in Spain”. In: *The Journal of Economic Inequality* 1.2, pp. 147–179. DOI: 10.1023/A:1026198429377.

— (2009). “The Invariance Properties of the Mutual Information Index of Multigroup Segregation”. In: *Research on Economic Inequality*. Ed. by Yves Flückiger, Sean F. Reardon, and Jacques Silber. Vol. 17. Emerald Group Publishing Limited, pp. 33–53. DOI: 10.1108/S1049-2585(2009)0000017005.

— (2011). “Entropy-Based Segregation Indices”. In: *Sociological Methodology* 41.1, pp. 159–194. DOI: 10.1111/j.1467-9531.2011.01237.x.

Müller, Walter and Markus Gangl, eds. (2003). *Transitions from Education to Work in Europe: The Integration of Youth into EU Labour Markets*. Oxford: Oxford University Press.

Müller, Walter and Yossi Shavit (1998). “The Institutional Embeddedness of the Stratification Process”. In: *From School to Work: A Comparative Study of Educational Qualifications and*

Occupational Destinations. Ed. by Walter Müller and Yossi Shavit. Oxford: Clarendon Press, pp. 1–48.

Müller, Walter, Susanne Steinmann, and Renate Ell (1998). “Education and Labour-Market Entry in Germany”. In: *From School to Work. A Comparative Study of Educational Qualifications and Occupational Destinations*. Ed. by Yossi Shavit and Walter Müller. Oxford: Clarendon Press, pp. 143–188.

Murnane, Richard J. and Sean F. Reardon (2018). “Long-Term Trends in Private School Enrollments by Family Income”. In: *AERA Open* 4.1, pp. 1–24. DOI: 10.1177/2332858417751355.

n.a. (1996). “The ‘Loi Jospin’. The Education Framework Act 1989 (Extracts)”. In: *Education in France. Continuity and Change in the Mitterrand Years, 1981–1995*. Ed. by Anne Corbett and Bob Moon. London and New York: Routledge, pp. 49–55.

OECD/Eurostat/UNESCO Institute for Statistics (2015). *ISCED 2011 Operational Manual: Guidelines for Classifying National Education Programmes and Related Qualifications*. Paris: OECD Publishing.

Orfield, Gary, Erica D. Frankenberg, and Chungmei Lee (2002). “The Resurgence of School Segregation”. In: *Educational Leadership* 60.4, pp. 16–20.

Orfield, Gary and John T. Yun (1999). *Resegregation in American Schools*. Cambridge, MA: The Civil Rights Project, Harvard University.

Owens, Ann (2016). “Inequality in Children’s Contexts: Income Segregation of Households with and without Children”. In: *American Sociological Review* 81.3, pp. 549–574. DOI: 10.1177/0003122416642430.

— (2017). “Racial Residential Segregation of School-Age Children and Adults: The Role of Schooling as a Segregating Force”. In: *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3.2, pp. 63–80. DOI: 10.7758/rsf.2017.3.2.03.

— (2018). “Income Segregation between School Districts and Inequality in Students’ Achievement”. In: *Sociology of Education* 91.1, pp. 1–27. DOI: 10.1177/0038040717741180.

Owens, Ann, Sean F. Reardon, and Christopher Jencks (2016). “Income Segregation Between Schools and School Districts”. In: *American Educational Research Journal* 53.4, pp. 1159–1197. DOI: 10.3102/0002831216652722.

Pendall, Rolf (2000). “Local Land Use Regulation and the Chain of Exclusion”. In: *Journal of the American Planning Association* 66.2, pp. 125–142. DOI: 10.1080/01944360008976094.

RDC of the Federal Statistical Office and Statistical Offices of the Länder (n.d.). *Census of 1970 and 1987 (Public Use Files)*.

Reardon, Sean F., Kendra Bischoff, Ann Owens, and Joseph B. Townsend (2018). “Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates”. In: *Demography* 55, pp. 2129–2160. DOI: 10.1007/s13524-018-0721-4.

Reardon, Sean F., Chad R. Farrell, Stephen A. Matthews, David O’Sullivan, Kendra Bischoff, and Glenn Firebaugh (2009). “Race and Space in the 1990s: Changes in the Geographic Scale of Racial Residential Segregation, 1990-2000”. In: *Social Science Research* 38.1, pp. 55–70. DOI: 10.1016/j.ssresearch.2008.10.002.

Reardon, Sean F. and Glenn Firebaugh (2002). “Measures of Multigroup Segregation”. In: *Sociological Methodology* 32.1, pp. 33–67. DOI: 10.1111/1467-9531.00110.

Reardon, Sean F., Elena Tej Grewal, Demetra Kalogrides, and Erica Greenberg (2012). “Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools”. In: *Journal of Policy Analysis and Management* 31.4, pp. 876–904. DOI: 10.1002/pam.21649.

Reardon, Sean F., Andrew D. Ho, B. R. Shear, E. M. Fahle, Demetra Kalogrides, H. Jang, B. Chavez, J. Buontempo, and R. DiSalvo (2019a). “Stanford Education Data Archive (Version 3)”. In: DOI: http://purl.stanford.edu/db586ns4974.

Reardon, Sean F., Demetra Kalogrides, and Kenneth Shores (2019b). “The Geography of Racial/Ethnic Test Score Gaps”. In: *American Journal of Sociology* 124.4, pp. 1164–1221. DOI: 10.1086/700678.

Reardon, Sean F. and Ann Owens (2014). “60 Years After Brown: Trends and Consequences of School Segregation”. In: *Annual Review of Sociology* 40.1, pp. 199–218. DOI: 10.1146/annurev-soc-071913-043152.

Reardon, Sean F. and John T. Yun (2001). “Suburban Racial Change and Suburban School Segregation, 1987-95”. In: *Sociology of Education* 74.2, pp. 79–101. DOI: 10.2307/2673164.

Reardon, Sean F., John T. Yun, and Tamela McNulty Eitle (2000). “The Changing Structure of School Segregation: Measurement and Evidence of Multiracial Metropolitan-Area School Segregation, 1989-1995”. In: *Demography* 37.3, pp. 351–364. DOI: 10.2307/2648047.

Reber, Sarah J. (2005). “Court-Ordered Desegregation: Successes and Failures Integrating American Schools since Brown versus Board of Education”. In: *Journal of Human Resources* 40.3, pp. 559–590. DOI: 10.3368/jhr.XL.3.559.

Reimers, Cordelia W. (1983). “Labor Market Discrimination Against Hispanic and Black Men”. In: *The Review of Economics and Statistics* 65.4, p. 570. DOI: 10.2307/1935925.

Renzulli, Linda A. and Lorraine Evans (2005). “School Choice, Charter Schools, and White Flight”. In: *Social Problems* 52.3, pp. 398–418. DOI: 10.1525/sp.2005.52.3.398.

Richards, Meredith P. (2014). "The Gerrymandering of School Attendance Zones and the Segregation of Public Schools: A Geospatial Analysis". In: *American Educational Research Journal* 51.6, pp. 1119–1157. DOI: 10.3102/0002831214553652.

Rossell, Christine H. (1975). "School Desegregation and White Flight". In: *Political Science Quarterly* 90.4, pp. 675–695. DOI: 10.2307/2148750.

— (1978). *Assessing the Unintended Impacts of Public Policy: School Desegregation and Resegregation.* Washington D.C.: National Institute of Education.

Rothwell, Jonathan T. and Douglas S. Massey (2010). "Density Zoning and Class Segregation in U.S. Metropolitan Areas: Density Zoning and Class Segregation". In: *Social Science Quarterly* 91.5, pp. 1123–1143. DOI: 10.1111/j.1540-6237.2010.00724.x.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek (2018). "IPUMS USA: Version 8.0". In: DOI: 10.18128/d010.v8.0.

Saporito, S. and D. Sohoni (2007). "Mapping Educational Inequality: Concentrations of Poverty among Poor and Minority Students in Public Schools". In: *Social Forces* 85.3, pp. 1227–1253. DOI: 10.1353/sof.2007.0055.

Saporito, Salvatore (2003). "Private Choices, Public Consequences: Magnet School Choice and Segregation by Race and Poverty". In: *Social Problems* 50.2, pp. 181–203. DOI: 10.1525/sp.2003.50.2.181.

— (2017a). "Irregularly-Shaped School Attendance Zones and Racial Integration". In: *Social Science Research* 64, pp. 299–318. DOI: 10.1016/j.ssresearch.2016.10.011.

— (2017b). "Shaping Income Segregation in Schools: The Role of School Attendance Zone Geography". In: *American Educational Research Journal* 54.6, pp. 1345–1377. DOI: 10.3102/0002831217724116.

Schafft, Kai A. (2006). "Poverty, Residential Mobility, and Student Transiency within a Rural New York School District". In: *Rural Sociology* 71.2, pp. 212–231. DOI: 10.1526/003601106777789710.

Shannon, Claude E. (1948). "A Mathematical Theory of Communication". In: *Bell System Technical Journal* 27.3, pp. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

Shapley, Lloyd. S. (1953). "A Value for N-Person Games". In: *Contributions to the Theory of Games.* Ed. by H. W. Kuhn and A. W. Tucker. Annals of Mathematical Studies 28. Princeton: Princeton University Press, pp. 307–317. DOI: 10.1515/9781400881970-018.

Shavit, Yossi and Walter Müller, eds. (1998). *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations.* Oxford: Clarendon Press.

Shorrocks, Anthony F. (2013). “Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value”. In: *The Journal of Economic Inequality* 11.1, pp. 99–126. DOI: 10.1007/s10888-011-9214-z.

Smock, Pamela J. and Franklin D. Wilson (1991). “Desegregation and the Stability of White Enrollments: A School-Level Analysis, 1968-84”. In: *Sociology of Education* 64.4, pp. 278–292. DOI: 10.2307/2112708.

Stroub, Kori J. and Meredith P. Richards (2013). “From Resegregation to Reintegration: Trends in the Racial/Ethnic Segregation of Public Schools, 1993–2009”. In: *American Educational Research Journal* 50.3, pp. 497–531. DOI: 10.3102/0002831213478462.

Strumbelj, Erik and Igor Kononenko (2010). “An Efficient Explanation of Individual Classifications Using Game Theory”. In: *Journal of Machine Learning Research* 11, pp. 1–18.

Taeuber, Karl E. and David R. James (1983). “Racial Segregation Among Public and Private Schools: A Response”. In: *Sociology of Education* 56.4, pp. 204–207. DOI: 10.2307/2112550.

Theil, Henri (1967). *Economics and Information Theory*. Amsterdam: North Holland.

— (1971). *Principles of Econometrics*. New York: Wiley.

— (1972). *Statistical Decomposition Analysis*. Amsterdam: North Holland.

Theil, Henri and Anthony J. Finizza (1971). “A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts”. In: *Journal of Mathematical Sociology* 1.2, pp. 187–193. DOI: 10.1080/0022250X.1971.9989795.

Thelen, Kathleen A. (2004). *How Institutions Evolve: The Political Economy of Skills in Germany, Britain, the United States, and Japan*. Cambridge: Cambridge University Press.

Watts, Martin (1998). “The Analysis of Sex Segregation: When Is Index Measurement Not Index Measurement?” In: *Demography* 35.4, pp. 505–508. DOI: 10.2307/3004018.

— (2015). “The Use and Abuse of Entropy-Based Segregation Indices”. In: *6th Meeting of the Society for the Study of Economic Inequality*. DOI: http://www.ecineq.org/ecineq_lux15/FILESx2015/CR2/p217.pdf.

Weeden, Kim A. (1998). “Revisiting Occupational Sex Segregation in the United States, 1910-1990: Results from a Log-Linear Approach”. In: *Demography* 35.4, pp. 475–487. DOI: 10.2307/3004015.

— (2004). “Profiles of Change: Sex Segregation in the United States, 1910-2000”. In: *Occupational Ghettos: The Worldwide Segregation of Women and Men*. Ed. by Maria Charles and David B. Grusky. Stanford, CA: Stanford University Press, pp. 131–78.

Welch, Finis and Andrew Light (1987). *New Evidence on School Desegregation.* Washington D.C.: U.S. Commission on Civil Rights.

Winkler, Richelle L. and Kenneth M. Johnson (2016). “Moving Toward Integration? Effects of Migration on Ethnoracial Segregation Across the Rural-Urban Continuum”. In: *Demography* 53.4, pp. 1027–1049. DOI: 10.1007/s13524-016-0479-5.

Winship, Christopher (1977). “A Revaluation of Indexes of Residential Segregation”. In: *Social Forces* 55.4, pp. 1058–1066. DOI: 10.2307/2577572.

Wolbers, M. H. J. (2007). “Patterns of Labour Market Entry: A Comparative Perspective on School-to-Work Transitions in 11 European Countries”. In: *Acta Sociologica* 50.3, pp. 189–210. DOI: 10.1177/0001699307080924.

Young, H. P. (1985). “Monotonic Solutions of Cooperative Games”. In: *International Journal of Game Theory* 14.2, pp. 65–72. DOI: 10.1007/BF01769885.

Appendix A: R packages `segregation` and `shapley`

The `segregation`[1] package includes functionality to calculate entropy-based segregation measures, namely the Mutual Information Index (M) and the Theil Index (H), which is the normalized version of the M index. The package also includes several methods for decomposing the index into between/within components and into local segregation scores, as well as methods to decompose differences in segregation indices. All of these methods have arguments to obtain standard errors and confidence intervals through bootstrapping.

The `shapley`[2] package provides a generic algorithm for Shapley value decompositions. By combining it with the `segregation` package, it can be used to easily specify custom decomposition problems.

A.1 Data format

For the examples, I will use a dataset built into the `segregation` package, `schools00`. This dataset contains data on 2,045 schools across 429 school districts in three U.S. states. For each school, the dataset records the number of Asian, Black, Hispanic, White, and Native American students. The `segregation` package requires data in long form (because most segregation data comes in this form), not in the form of contingency tables. Hence, each row of the `schools00` dataset is a unique combination of a given school and a racial group, and the column `n` records the number of students for this combination:

```
library("segregation")
options(digits = 3)

head(schools00[, c("school", "race", "n")])
```

[1] `https://elbersb.github.io/segregation`
[2] `https://github.com/elbersb/shapley`

```
#>    school  race   n
#> 1:   A1_1 asian   2
#> 2:   A1_1 black  14
#> 3:   A1_1  hisp  30
#> 4:   A1_1 white 351
#> 5:   A1_2 black   9
#> 6:   A1_2  hisp 101
```

Note that in the first school, `A1_1`, there are no Native American students. Hence, that row is missing.

If the data is in the form of contingency tables, it is possible to use `matrix_to_long()` to convert them to the long format required for the package. As an example:

```
(m = matrix(c(10, 20, 30, 30, 20, 10), nrow = 3))
#>      [,1] [,2]
#> [1,]   10   30
#> [2,]   20   20
#> [3,]   30   10
colnames(m) <- c("Black", "White")
matrix_to_long(m, group = "race", unit = "school")
#>    school  race  n
#> 1:      1 Black 10
#> 2:      2 Black 20
#> 3:      3 Black 30
#> 4:      1 White 30
#> 5:      2 White 20
#> 6:      3 White 10
```

The `group` and `unit` arguments are optional.

A.2 Computing the M and H indices

Compute the M and H indices using `mutual_total()`:

```
mutual_total(schools00, "race", "school", weight = "n")
#>    stat   est
#> 1:    M 0.426
#> 2:    H 0.419
```

Interpretation of a single M value is not straightforward, because it is not normalized. However, the H can range from 0 to 1, so a value of 0.419 would indicate moderate segregation.

The second argument to `mutual_total()` refers to the groups, while the third argument refers to the units. Switching groups and units does not affect the M index, but does change the H index:

```
mutual_total(schools00, "school", "race", weight = "n")
#>    stat    est
#> 1:    M 0.4255
#> 2:    H 0.0564
```

This is because the `segregation` package always divides by the marginal *group* entropy, and it would here hence divide by the entropy of the school distribution, which we would expect to be much larger (as there are many more schools than racial groups). To check, we can use the `entropy()` function:

```
(entropy(schools00, "race", weight = "n"))
#> [1] 1.02
(entropy(schools00, "school", weight = "n"))
#> [1] 7.54
```

Therefore, if the H index is used, it is important to specify the groups and units correctly.

For inference (discussed in more detail below), it is possible to use bootstrapping to obtain standard errors and confidence intervals:

```
mutual_total(schools00, "race", "school", weight = "n",
             se = TRUE, CI = .95, n_bootstrap = 500)
#> 500 bootstrap iterations on 877739 observations
#>    stat   est       se          CI    bias
#> 1:    M 0.422 0.000774 0.420,0.424 0.00358
#> 2:    H 0.415 0.000705 0.414,0.417 0.00351
```

As there a large number of observations, the standard errors are very small.

A.3 Between-Within decomposition

We might wonder whether segregation is different across the three different states. We can compute their segregation indices manually (just showing the M for simplicity):

```
split_schools <- split(schools00, schools00$state)
mutual_total(split_schools$A, "race", "school", weight = "n")[1, ]
#>    stat   est
#> 1:    M 0.409
mutual_total(split_schools$B, "race", "school", weight = "n")[1, ]
#>    stat   est
#> 1:    M 0.255
mutual_total(split_schools$C, "race", "school", weight = "n")[1, ]
#>    stat   est
#> 1:    M 0.345
```

Clearly, state A is more segregated than state C, which in turn shows higher school segregation than B. One of the advantages of entropy-based segregation indices is that these three state-specific indices have a simple relationship to the overall index. This is just the between/within decomposition: Total segregation can be decomposed into a term that measures how much the distribution of racial groups differs *between* states, and into a term that measures segregation *within* states. If we have S states (or "super-units" more generally), where each school belongs to exactly one state, the M index can be decomposed as follows:

$$M(\mathbf{T}) = M(\mathbf{S}) + \sum_{s=1}^{S} p_s M(\mathbf{T}_s),$$

where $\mathbf{T}$ is the full $U \times G$ contingency table, $\mathbf{S}$ is the aggregated contingency table of dimension $S \times G$, p_s is the population proportion of state s (such that $\sum_{s=1}^{S} p_s = 1$), and $\mathbf{T}_s$ is the subset of rows of $\mathbf{T}$ belonging to state s. Put in simple terms, the M index can be decomposed into a between-state segregation index, plus a weighted average of within-state M indices.

For the H index, we are dividing the above formula by $E(\mathbf{T})$. This makes the formula a bit more complicated, because the normalization has to be offset in the decomposition:

$$H(\mathbf{T}) = H(\mathbf{S}) + \sum_{s=1}^{S} \frac{E(\mathbf{T}_s)}{E(\mathbf{T})} p_s H(\mathbf{T}_s),$$

where $E(\cdot)$ is again the entropy of the marginal group distribution. Note that $E(\mathbf{T}) = E(\mathbf{S})$, because the group marginal distributions are identical.

To compute the decomposition using the `segregation` package, use:

```
# total segregation
(total <- mutual_total(schools00, "race", "school", weight = "n"))
#>    stat   est
#> 1:    M 0.426
#> 2:    H 0.419

# between-state segregation:
#     how much does the racial distributions differ across states?
(between <- mutual_total(schools00, "race", "state", weight = "n"))
#>    stat    est
#> 1:    M 0.0992
#> 2:    H 0.0977

# within-state segregation:
#     how much segregation is there within states?
(mutual_total(schools00, "race", "school", within = "state", weight = "n"))
#>    stat   est
#> 1:    M 0.326
#> 2:    H 0.321
```

Note that $0.426 = 0.0992 + 0.326$ and $0.419 = 0.0977 + 0.321$. The results indicate that about 75% of the segregation is within states. In other words, differences in the racial composition of the three different states account for less than 25% of segregation.

By using `mutual_total()` with the `within` argument, we can obtain the overall within component, but we do not obtain the decomposition by state. To do so, we can use `mutual_within()`:

```
(within <- mutual_within(schools00, "race", "school",
                         within = "state", weight = "n", wide = TRUE))
#>    state     M     p     H ent_ratio
#> 1:     A 0.409 0.277 0.497     0.809
#> 2:     B 0.255 0.404 0.268     0.936
```

```
#> 3:      C 0.345 0.320 0.361     0.940
```

This is a much simpler way to obtain state-specific segregation scores compared to subsetting manually, as shown in the beginning of this section. In addition to the M and H indices, we also obtain `p`, the population proportion of the state (p_s above), and `ent_ratio`, which is $E(\mathbf{T}_s)/E(\mathbf{T})$ from above. Hence, one can recover the total within-component using

```
with(within, sum(M * p))
#> [1] 0.326
with(within, sum(H * p * ent_ratio))
#> [1] 0.321
```

which is exactly the same as before. The quantity $p_s M(\mathbf{T}_s)$ is itself of interest, because it shows how much the states contribute to the segregation total, when taking into account their size. By adding the between component, we can calculate the contribution of the four components:

```
# merge into a vector
components <- c(between$est[1], within$M * within$p)
names(components) <- c("Between", "A", "B", "C")
signif(100 * components / total$est[1], 3)
#> Between       A       B       C
#>    23.3    26.6    24.2    25.9
```

Each of the four components contributes about a quarter to the total segregation of 0.426. Note that state A is the smallest state (27.7% of the population), but contributes the largest percentage (26.6%) to total segregation. Hence, the decomposition shows that it is important to look at both p_s, the state sizes, as well as $M(\mathbf{T}_s)$, within-state segregation.

A.4 Local segregation

The M index (but not the H index) allows for another decomposition, into local segregation scores. To define this decomposition, let $p_{g|u} = t_{ug}/t_{u\cdot}$ be the conditional probability of being in group g, given that one is in unit u. We can then define the *local segregation score of unit u* as

$$L_u = \sum_{g=1}^{G} p_{g|u} \log \frac{p_{g|u}}{p_{\cdot g}}.$$

The weighted average of the L_u is again $M(\mathbf{T})$, i.e. $M(\mathbf{T}) = \sum_{u=1}^{U} p_{u\cdot} L_u$.

To obtain the local segregation scores L_u, along with the marginal weights $p_{u\cdot}$, use `mutual_local()`:

```
mutual_local(schools00, "race", "school", weight = "n", wide = TRUE)
#>       school    ls        p
#>    1:   A1_1 0.183 0.000452
#>    2:   A1_2 0.183 0.000498
#>    3:   A1_3 0.276 0.000664
#>    4:   A1_4 0.137 0.000569
#>    5:   A2_1 0.359 0.000426
#>   ---
#> 2041: C165_1 0.317 0.000457
#> 2042: C165_2 0.384 0.000530
#> 2043: C165_3 0.297 0.000565
#> 2044: C166_1 0.307 0.001159
#> 2045: C167_1 0.317 0.000535
```

Local segregation scores are based on much less data than the full M index, so it often makes sense to obtain confidence intervals. The following code computes the correlation between the length of the 95% confidence interval in relation to the size of each school:

```
localse <- mutual_local(schools00, "race", "school", weight = "n",
                        se = TRUE, wide = TRUE, n_bootstrap = 1000)
localse$lengthCI <- sapply(localse$ls_CI, diff)
with(localse, cor(p, lengthCI))
#> [1] -0.317
```

Although the relationship is far from deterministic, larger schools have shorter confidence intervals.

Because the M is symmetric, local segregation scores can also be obtained for groups. The equivalent definition for the *local segregation score of group* g is then

$$L_g = \sum_{u=1}^{U} p_{u|g} \log \frac{p_{u|g}}{p_{u\cdot}},$$

and, as expected, $M(\mathbf{T}) = \sum_{g=1}^{G} p_{\cdot g} L_g$.

To obtain these scores, switch the group and unit arguments in `mutual_local`:

```
(localg <- mutual_local(schools00, "school", "race", weight = "n", wide = TRUE))
#>      race    ls       p
#> 1:  asian 0.629 0.02255
#> 2:  black 0.881 0.19015
#> 3:   hisp 0.777 0.15170
#> 4:  white 0.184 0.62809
#> 5: native 1.434 0.00751
```

These results show that the racial groups experience very different levels of segregation: White students are less segregated than Asian, Black, Hispanic, and, especially, Native American students.

A.5 Inference

The four main functions of the packages, `mutual_total()`, `mutual_within()`, `mutual_local()`, and `mutual_difference()` all support inference through bootstrapping. Inference for segregation indices is tricky, and the standard error estimates and confidence intervals should not be trusted too much when there is little data, and especially when the segregation index is very close to either 0 or maximum segregation.

To estimate standard errors and confidence intervals, use `se = TRUE`. The coverage of the confidence interval can be specified in the `CI` argument. The number of bootstrap iterations can be specified as well:

```
(se <- mutual_total(schools00, "race", "school", weight = "n",
                    se = TRUE, CI = .95, n_bootstrap = 500))
#> 500 bootstrap iterations on 877739 observations
#>    stat   est       se          CI    bias
#> 1:    M 0.422 0.000708 0.420,0.423 0.00363
#> 2:    H 0.415 0.000655 0.414,0.417 0.00355
```

The confidence intervals are based on the percentiles from the bootstrap distribution, and hence require a large number of bootstrap iterations for valid interpretation. The estimate `est` that is reported in the results has already been "debiased," i.e. the bias that has been estimated from the bootstrap distribution (which is reported in `bias`) has been subtracted from the usual maximum-likelihood estimate that we would obtain from `mutual_total` with `se = FALSE`. The confidence interval is centered around the debiased estimate.

On balance, confidence intervals are preferred over the standard error because the bootstrap distribution can be skewed, especially when segregation is very low or very high. For this example, we can see that the standard errors provide almost identical coverage to the confidence intervals, as

```
# M
with(se, c(est[1] - 1.96 * se[1], est[1] + 1.96 * se[1]))
#> [1] 0.421 0.423
# H
with(se, c(est[2] - 1.96 * se[2], est[2] + 1.96 * se[2]))
#> [1] 0.414 0.417
```

provide effectively the same coverage as the confidence intervals obtained from the percentile bootstrap.

Whenever the bootstrap is used, the bootstrap distributions for each parameter are reported in an attribute `bootstrap` of the returned object. This can be used, for instance, to check whether the bootstrap distribution is skewed. The following code computes local segregation scores for all schools, and then computes the quantiles of the centered bootstrap distribution for school `C137_9`, which has a very low local segregation score:

```
local <- mutual_local(schools00, "race", "school", weight = "n",
                      se = TRUE, CI = .95, n_bootstrap = 500)
# pick bootstrap distribution of local segregation scores for school C137_9
boot <- attr(local, "bootstrap")
ls_school <- boot[school == "C137_9" & stat == "ls", boot_est]
quantile(ls_school - mean(ls_school), probs = c(0.01, 0.25, 0.5, 0.75, 0.99))
#>         1%       25%       50%       75%       99%
#> -0.006522 -0.002744 -0.000521  0.002286  0.010804
```

For this school, the bootstrap distribution is skewed. If precise inference about this specific school is needed, the standard error should not be interpreted, and the confidence interval should only be interpreted when the number of bootstrap iterations is large.

A.6 Decomposing differences in indices using the marginal-structural method

The command `mutual_difference()` can be used to decompose differences in segregation, as described in the second chapter. The default, and recommended method, is to use `method = shapley` (or `method = shapley_detailed`, which adds information about the decomposition of local segregation scores). The other methods (`mrc`, `km`) exist mostly for testing purposes, and are not recommended. They can, however, be used to replicate the findings of chapter 2.

```
mutual_difference(schools00, schools05, "race", "school", weight = "n")
#>                stat      est
#> 1:               M1  0.42554
#> 2:               M2  0.41339
#> 3:             diff -0.01215
#> 4:        additions -0.00341
#> 5:         removals -0.01141
#> 6: group_marginal  0.01855
#> 7:  unit_marginal -0.01239
#> 8:     structural -0.00349
```

In this case, structural change plays no role in explaining segregation change.

This method also supports inference by setting `se = TRUE`.

A.7 Decomposing differences in indices using custom Shapley decompositions

Using the `shapley` package to undertake custom decompositions first requires setting up a value function. This value function takes as parameters the factors that are involved in computing the counterfactual. For this example, I will set up a value function that decomposes the change in segregation between the `schools00` (year 2000) and `schools05` (year 2005) datasets by racial group. As there are five racial groups, the decomposition involves five factors. The value function works as follows: If one of the factors is supplied, it will take the racial group counts from 2005. When the factor is absent, the value function will take the racial group counts from 2000. Hence, when all factors are absent, the value function returns the segregation observed in 2000, and when all factors are supplied, the value function returns the segregation observed in 2005. All other factor combinations are counterfactuals.

To start, I first combine the two datasets into one, dropping schools that are only present at either time point.

```
schools <- intersect(schools00$school, schools05$school)
merged <- merge(schools00, schools05, all = TRUE,
    by = c("state", "district", "school", "race"),
    suffixes = c("2000", "2005"))
merged <- merged[merged$school %in% schools, ]
merged$n2000 <- with(merged, ifelse(is.na(n2000), 0, n2000))
merged$n2005 <- with(merged, ifelse(is.na(n2005), 0, n2005))
head(merged)
#> state district school   race n2000 n2005
#>     A       A1   A1_3  asian     0     3
#>     A       A1   A1_3  black    10    13
#>     A       A1   A1_3   hisp   217   412
#>     A       A1   A1_3  white   356   817
#>     A       A1   A1_3 native     0     2
#>     A       A2   A2_2  asian     0     2
```

The value function is then constructed as follows:

```
vfun <- function(factors) {
    # use racial group counts from either 2000 or 2005
    merged$counterf <- with(merged, ifelse(race %in% factors, n2005, n2000))
    # compute H index
    mutual_total(merged, "race", "school", weight = "counterf")[, est][2]
}
```

We can check whether the value function works as expected by comparing to the observed values in 2000 and 2005:

```
# segregation in 2000
vfun(c())
#> [1] 0.412
mutual_total(merged, "race", "school", weight = "n2000")[, est][2]
#> [1] 0.412
```

```r
# segregation in 2005
vfun(c("asian", "black", "hisp", "white", "native"))
#> [1] 0.39
mutual_total(merged, "race", "school", weight = "n2005")[, est][2]
#> [1] 0.39
```

The remaining step is to call the `shapley` function:

```r
library("shapley")
shapley(vfun, c("asian", "black", "hisp", "white", "native"))
#>    factor    value
#> 1:  asian -0.00678
#> 2:  black -0.01751
#> 3:   hisp -0.02028
#> 4:  white  0.02403
#> 5: native -0.00100
```

The observed change in segregation, `0.419 - 0.389 = 0.03`, is thus decomposed into five factors. What jumps out in this example is that the enrollments counts for Black and Hispanic students changed in ways that reduced segregation, while the enrollment counts of White students increased segregation.

When the number of factors grows larger, the `shapley` will become to slow to compute the closed-form solution. In this case, the `shapley_sampled` function can be used (for five factors, this will usually be slower):

```r
shapley_sampled(vfun, c("asian", "black", "hisp", "white", "native"))
#>    factor     value value_se iterations
#> 1:  asian -0.006777 7.13e-05        104
#> 2:  black -0.017532 4.37e-05        101
#> 3:   hisp -0.020237 7.12e-05        101
#> 4:  white  0.024130 8.26e-05        137
#> 5: native -0.000973 1.01e-05        101
```

The sampled values align very closely with the closed-form solutions.