

The Springer Series on Demographic Methods
and Population Analysis 42

Mark Fossett

New Methods for Measuring and Analyzing Segregation



Springer Open

The Springer Series on Demographic Methods and Population Analysis

Volume 42

Series Editor

Kenneth C. Land, Duke University

In recent decades, there has been a rapid development of demographic models and methods and an explosive growth in the range of applications of population analysis. This series seeks to provide a publication outlet both for high-quality textual and expository books on modern techniques of demographic analysis and for works that present exemplary applications of such techniques to various aspects of population analysis.

Topics appropriate for the series include:

- General demographic methods
- Techniques of standardization
- Life table models and methods
- Multistate and multiregional life tables, analyses and projections
- Demographic aspects of biostatistics and epidemiology
- Stable population theory and its extensions
- Methods of indirect estimation
- Stochastic population models
- Event history analysis, duration analysis, and hazard regression models
- Demographic projection methods and population forecasts
- Techniques of applied demographic analysis, regional and local population estimates and projections
- Methods of estimation and projection for business and health care applications
- Methods and estimates for unique populations such as schools and students

Volumes in the series are of interest to researchers, professionals, and students in demography, sociology, economics, statistics, geography and regional science, public health and health care management, epidemiology, biostatistics, actuarial science, business, and related fields.

More information about this series at <http://www.springer.com/series/6449>

Mark Fossett

New Methods for Measuring and Analyzing Segregation

 Springer Open

Mark Fossett
Department of Sociology
Texas A&M University
College Station, TX, USA



ISSN 1389-6784 ISSN 2215-1990 (electronic)
The Springer Series on Demographic Methods and Population Analysis
ISBN 978-3-319-41302-0 ISBN 978-3-319-41304-4 (eBook)
DOI 10.1007/978-3-319-41304-4

Library of Congress Control Number: 2017930612

© The Editor(s) (if applicable) and the Author(s) 2017. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution-NonCommercial 2.5 International License (<http://creativecommons.org/licenses/by-nc/2.5/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

This work is subject to copyright. All commercial rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on micro-films or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Abstract

In this monograph, I place indices used to measure the uneven distribution dimension of residential segregation in a new framework; I cast them as simple differences of group means on individual-level residential outcomes scored from area racial composition. The “difference-of-group-means” framework places all popular indices in a common measurement framework in which index scores are additively determined by individual residential attainments. This yields new and appealing options regarding substantive interpretations of the scores of segregation indices. It also brings important methodological benefits by creating the new possibility of joining the investigation of aggregate segregation and the investigation of individual-level residential attainments together in a single analysis. Specifically, segregation index scores now can be equated with the effect of group membership (e.g., race) on individual residential attainments, and thus variation in segregation over time and across cities can be equated to the ways that the effect of group membership varies over time and with city characteristics in multilevel models of individual residential attainments. Framing segregation indices in the difference-of-group-means framework has several other desirable consequences for segregation analysis. It creates opportunities to investigate segregation in new ways by permitting researchers to assess the impact of group membership on residential outcomes in the context of multivariate attainment models that if desired can include controls for other individual characteristics (e.g., language, education, income). Relatedly, it suggests a new basis on which to evaluate and compare segregation indices – whether the individual-level residential outcomes they register and reflect are relevant for theories of residential dynamics and/or are relevant for concerns about racial differences in socioeconomic attainments and life chances. Finally, the difference-of-group-means framework paves the way for developing refined versions of indices that are free of potentially problematic upward bias intrinsic to standard formulations of these indices. Significantly, adopting the new framework outlined here does not require breaking with previous conceptions of segregation; results of empirical analyses of segregation using traditional computing formulas can be exactly replicated within this framework even as several new options for measurement and analysis become available.

Preface

In this monograph, I review findings and observations I have accumulated while grappling with issues in segregation measurement over the past decade. My explorations in this area were motivated by three concerns. The first was that, while it is obvious to all concerned that residential segregation can potentially have important consequences for group differences in residential outcomes, the literature on segregation measurement does not provide formulations of segregation indices that make it clear exactly what implications index scores have for group differences in residential outcomes. In this regard, the measurement and analysis of segregation is on a different conceptual footing from standard approaches to measuring and analyzing intergroup disparity and inequality on other socioeconomic and stratification outcomes such as education, occupation, and income. Researchers investigating disparities in these other areas routinely assess inequality and disparities based on comparisons of group means on individual-level outcomes. Consequently, the connections between scores of measures of aggregate inequality have clear and direct implications for group differences in the attainments of individuals. In contrast, the literature on segregation measurement has not established how segregation index scores are connected to group differences on residential outcomes for individuals. This is surprising and unfortunate because the substantive relevance of segregation indices ultimately rests on the presumption that their scores carry important implications for group differences on individual residential outcomes and yet these implications have remained obscure. I address this concern here by introducing new formulations of popular segregation indices that place them in an overarching “difference-of-group-means” framework that clarifies exactly how segregation index scores are connected to group differences in individual-level residential outcomes.

The second concern motivating me was that the literature on segregation measurement and analysis did not provide a straightforward means for directly linking quantitative findings from studies of micro-level processes of residential attainment to findings for segregation index scores at the aggregate level (e.g., city-level segregation scores). As a result, the research literature has been divided into two important but largely disconnected traditions. One is a tradition of macro-level studies

that use aggregate-level index scores for cities to investigate how segregation varies across cities and over time; the other is a tradition of micro-level studies that examine how various individual-level residential attainments are related to social and economic characteristics of individuals and households such as income, education, nativity, English language ability, family type, and other related individual-level variables. The current state of the literature leaves researchers in both traditions in the frustrating situation of being unable to directly connect segregation index scores at the aggregate-level to the individual-level outcomes that are examined and modeled in micro-level residential attainment analyses. I address this concern here by drawing on the difference-of-group-means measurement framework to develop methods for linking index scores to individual-level residential attainment processes. In this new approach, segregation index scores now can be interpreted as the effect of group membership (e.g., race) on segregation-determining residential outcomes in an individual-level attainment model. The level of segregation in a city thus can now be assessed by estimating the effect of group membership on individual-level residential attainment in bivariate attainment model. More importantly, the model can be extended to a multivariate specification to properly take account of the role that nonracial characteristics (e.g., income) may play in shaping the level of segregation in a city. And the model can be further extended to multi-level specifications to take account of how city-level factors impact segregation net of the role of nonracial individual characteristics. Significantly, past findings of aggregate-level analyses can be exactly replicated and subsumed under this approach while giving researchers many new options for analysis.

The third concern motivating me was that, under the current state of segregation measurement, many interesting and important research questions cannot be addressed because segregation index scores exhibit problematic behavior under a wide range of commonly occurring conditions. In particular, all indices of uneven distribution are subject to inherent positive bias that can render their scores untrustworthy and potentially misleading in a variety of situations – for example, when segregation is measured at small spatial scales (e.g., at the block level) or when the groups involved in the segregation comparison are small and/or are imbalanced in size. This presents severe obstacles to many interesting and important lines of inquiry in segregation research. For example, it precludes quantitative study of segregation for newly arriving immigrant and migrant groups because, by definition, the groups initially are small in both absolute size and relative size in comparison to established population groups. Similarly, it precludes study of segregation among narrowly defined subgroups with a population (e.g., foreign- and native-born Latinos, high-income Whites and Blacks, etc.) because one or both subgroups often are small in absolute and/or relative size. Additionally, it potentially frustrates investigation of segregation dynamics using agent simulation models because studies in this tradition routinely examine segregation at small spatial scales.

The impact of these concerns on segregation research is substantial, pervasive, and hard to overstate. It has led researchers to routinely adopt two “defensive” practices. One practice is to use various ad hoc guidelines to “screen” cases to avoid measuring segregation in situations where index scores cannot be trusted. The other

practice is to differentially weight cases to minimize the undesirable impact of bias on index scores even after cases have been “screened” to eliminate those where index scores are most problematic. The first practice prevents researchers from undertaking many studies that otherwise would be conducted and thus sharply restricts the scope of segregation studies. In addition, it draws on ad hoc guidelines that at best are crude and at worst have uncertain effectiveness. The second practice of differentially weighting cases is predicated on the implicit recognition that the first practice of screening cases cannot adequately deal with the problem of bias. Unfortunately, differential weighting of cases is itself inadequate. First and foremost, it leaves index scores untrustworthy on a case-by-case basis and so one cannot discuss and compare cases – otherwise weighting would be unnecessary. Second, while the strategy permits researchers to avoid “draconian” screening of cases and thus larger nominal sample sizes, differential weighting in the end amounts to assessing segregation patterns and trends based on the small subset of cases that get large weights.

I address this unsatisfying state of affairs by developing and introducing refined versions of popular segregation indices that provide trustworthy measurements of segregation over a much broader range of situations than standard measures. I demonstrate that the resulting unbiased measures have attractive properties and provide researchers the previously unavailable option of dealing with index bias directly at the point of measurement on a case-by-case basis.

As I worked to address the three concerns just mentioned, I increasingly took interest in a fourth concern – the question of whether different segregation indices yielded similar or different results and, if different, under what conditions and why. Conventional wisdom in the segregation measurement literature has been that the most widely used measures of uneven distribution tend to give similar results. But I found discrepancies between indices were common when I measured segregation over broader samples of cases and group comparisons. At first I thought the large discrepancies between scores of different indices might be a by-product of the problem of index bias. After all, using broader samples tends to include cases that are more susceptible to being adversely affected by the problem of bias, and previous methodological studies had reported that indices vary in susceptibility to scores being inflated by index bias. But on investigating the issue further, I found that the role of bias was only a minor part of the story as discrepancies between scores for different indices persisted even when using refined versions of the indices that were free of the influence of bias.

The difference-of-group-means framework provided a useful perspective for exploring this issue and led me to recognize that the discrepant scores I observed reflected an aspect of uneven distribution that is not generally widely appreciated, namely, index sensitivity, or lack thereof, to whether displacement from even distribution is concentrated or dispersed. My goal in exploring this issue was different in nature from my goals in addressing the first three concerns I noted. In this case, I was not seeking to make progress toward solving technical problems in measuring and analyzing segregation. Instead, my goal was to clarify the nature of the differences between indices to better account for why different indices sometimes yield

different results. In the end, I concluded the issue could be framed succinctly in terms of index sensitivity to whether group displacement from even distribution is concentrated and dispersed. At any given nontrivial level of group displacement from even distribution, groups can be concentrated in a way that produces homogeneous areas for both groups, or groups can be dispersed in a way that minimizes homogeneous areas. Indices vary in their sensitivity to this aspect of uneven distribution. For example, the widely used index of dissimilarity (D) takes the same value regardless of whether displacement is concentrated or dispersed, while the separation index (S) takes higher values when displacement is concentrated and takes low values when displacement is widely dispersed.

I am hardly the first to recognize the technical basis for this potential difference between indices. But I believe my discussion and review of these issues makes useful new contributions to the literature on segregation measurement. First off, the analyses I report here document that important discrepancies between different index scores are much more common than previous methodological studies have suggested. Second, the difference-of-means framework for measuring segregation I introduce here provides a new basis for understanding exactly how different indices can yield discrepant index scores. Finally, I offer analytic exercises and empirical case studies to further clarify the basis of differences between indices and dispel certain misconceptions regarding of these issues.

My hope is that this monograph will contribute to a better understanding of the issues examined here and also will provide useful practical strategies for measuring and analyzing segregation. Looking back on the decade of work reflected here, I can see with hindsight that the core issues are closely interconnected. Establishing how segregation index scores related to group differences in residential outcomes was a necessary step for developing methods for conducting micro-level analysis of individual-level residential attainments that could directly account for overall segregation in a city at the aggregate level. Discovering that the residential attainments in question were rooted in a simple construct – the pairwise group proportions in the area of residence – then paved the way for a further discovery, namely, that troublesome problem of index bias could be eliminated by making surprisingly simple refinements in the calculation of pairwise group proportions. Thinking more carefully about the individual-level residential outcomes that are registered by different indices led to a better understanding of the differences between concentrated and dispersed displacement from even distribution.

The interconnections among the issues are clearer in hindsight. If I had recognized them from the start, I would have avoided muddling around for so long. I offer my findings and observations on these and related matters here in hopes that others will find them useful. I apologize in advance for the many limitations of this study but also suggest that it occasionally offers original insights and new options for segregation measurement and analysis that I hope can help other researchers move the study of segregation forward.

Many organizations and many people have provided support and encouragement that helped make my work possible. Over the past decade, I was fortunate to receive funding support for projects that helped me develop findings and observations

included in this monograph. They included National Institutes of Health research grants R43HD038199 and R44HD038199 “Simulating Residential Segregation Dynamics: Phases I & II”; a proposal development grant from the Mexican American and Latino Research Center at Texas A&M University, College Station; and National Science Foundation research grant SES 1024390 “New Methods for Segregation Research.” Of course the funding agencies are not responsible for and do not necessarily endorse the findings and conclusions I offer. Finally, I also acknowledge a faculty development leave from Texas A&M University that was crucial for completing the first full draft of the monograph. I also thank the College of Liberal Arts and the Open Access to Knowledge (OAK) program at Texas A&M University which have provided generous funding to help publish this monograph as an open access work.

I have always received encouragement and support from my colleagues and good friends in the Sociology Department at Texas A&M University. In particular, I must mention Jane Sell, Dudley Poston, and Rogelio Saenz to offer special thanks for their support over this period of extended effort. I thank Wenquan “Charles” Zhang for engaging me in many stimulating and productive discussions on the issues addressed in this study and for collaborating with me as co-investigator on the aforementioned NSF project and associated empirical analyses that draw on the new measures and methods introduced in this work. Warner Henson III, currently a doctoral student in the Sociology Department at Stanford University, deserves special acknowledgment for helping me establish the difference-of-means formulation of the Theil index while an undergraduate major in sociology here at Texas A&M University. I also must thank Amber Fox Crowell, at the time a doctoral student in sociology at Texas A&M University and now an assistant professor at California State University-Fresno, who served as research assistant on research projects in which the measures and methods introduced here were developed and refined. Her insights, questions, and suggestions have been extremely helpful. She has a deep grasp of the potential value the measures and methods can bring to empirical analyses and has applied them with great success in her dissertation project and her projects as a postdoctoral researcher. I also offer a note of appreciation to the many students who have attended informal workshops on segregation measurement and analysis on a regular basis over the past several years including, in rough chronological order, Warren Waren, Lindsay Howden, Amber Fox Crowell, Gabriel Amaro, Bianca Manago, Jennifer Davis, Jessica Barron, Nicole Jones, Bo Hee Yoon, Brittany Rico, Melissa Sanchez, Chiying Huang, Xuanren Wang, Katelyn Polk, Bridget Clark, Danielle Deng, Nathanael Rosenheim, Cassidy Castiglione, and Xinyuan Zou. Their questions, puzzlements, and suggestions helped me develop better ways of explaining and thinking about some of the material presented here. As ever throughout my career, I benefit from the voice of my dissertation supervisor, mentor, and friend, Omer Galle, which always present in the back of my mind encouraging me and challenging me.

I also must acknowledge the amazing support Ms. Evelien Bakker and Ms. Bernadette Deelen-Mans with Springer have given during the process of bringing this monograph to completion. Apparently, it is impossible to exhaust their patience

and goodwill. They have been encouraging, helpful, and pleasant at every step in the process. For that I am truly grateful.

I close by thanking my wife, Betsy, and our children Kate, Tyler, and Lane for their support and patience. They have brought, and continue to bring, great joy to my life. They have kept me grounded through the ups and downs of the too-long gestation period for this monograph. I love them more than they know and appreciate that they put up with me and indulge my passion for the study of social inequality and segregation.

College Station, TX, USA
December 2015
July 2017

Mark Fossett