Segregated from the Start

Methodology

Erica Greenberg and Tomas Monarrez

September 2019

Data for the figures presented in this feature come from three sources measuring school and program enrollment by race or ethnicity nationally. The first is the National Survey of Early Care and Education, a nationally representative study of center- and home-based early childhood programs conducted in 2012. The study draws on integrated surveys of providers and households to offer comprehensive information on early childhood program use and availability. Our analyses leverage public-use center- and home-based provider files and incorporate sampling weights and related information to support nationally representative estimates. We restrict all analyses to programs serving at least five children to ensure the possibility of integration.

To supplement these data, we use additional data from the Common Core of Data and the Private School Universe Survey, both from the National Center for Education Statistics. These data describe the universe of schools serving grades K–12 for students attending regular schools in the US mainland states during the 2011–12 school year. The Common Core of Data is a census, providing enrollment data for every public school in the country. The 2011–12 Private School Universe Survey includes data on 91.8 percent of the country's private schools, as reported in the survey's general documentation.

Segregation indexes summarize the distribution of racial or ethnic groups among units (in our case, programs) within a system (the country). The first figure presents a nonparametric summary of this distribution, showing estimates of the density of programs within 10 bins of the black or Hispanic program population share separately for early childhood education and for kindergarten and first grade. In the rest of the figures, we summarize segregation in this distribution using a single index, the dissimilarity index, computed as follows:

$$Dissimilarity = \sum_i \frac{p_i|m_i - M|}{2PM(1 - M)}.$$

Here, p_i is total enrollment in program $i = 1, \dots, N$ and N is the total number of programs in the country; m_i is the black or Hispanic share of program students; M is the black or Hispanic share of students nationally; and P is the total number of students in the country.

Intuitively, the dissimilarity index measures racial or ethnic imbalance between programs. It can be interpreted as the share of black or Hispanic students that would have to switch programs to achieve complete integration, as a ratio of the same share, but starting from a completely segregated status quo. High values of dissimilarity mean that a large share of students would have to switch programs to achieve integration. In a perfectly integrated environment, the index equals 0, and in a perfectly segregated system, the index equals 1.

Notably, this formulation of segregation imposes the assumption that the national black or Hispanic share is the racial or ethnic composition of a racially balanced program. Nevertheless, the results presented in this feature are robust to the use of other segregation indexes, including the isolation index and the variance ratio index. Isolation indexes are based on exposure to diversity, summarizing the average racial or ethnic composition of the programs black or Hispanic students attend. Variance ratio indexes are based on isolation measures but include an adjustment for overall racial or ethnic composition.

The indexes presented in this feature are measured on a national scale. In that sense, these indexes assume the entire country is a school system to be evaluated based on its segregation. This contrasts with more common approaches to measuring segregation that focus on specific geographies, be they school districts, metropolitan areas, or states. Intuitively, the country has a higher scope for segregation than a small community because it is larger, more populous, and more racially and ethnically diverse than the average community. This should be considered when interpreting our findings, as it implies that our estimates may change at other levels of geography.

About the Authors

Erica Greenberg is a senior research associate in the Center on Education Data and Policy at the Urban Institute.

Tomas Monarrez is a research associate in the Center on Education Data and Policy.

Acknowledgments

For more information on this project, see Tomas Monarrez, "Segregated Neighborhoods, Segregated Schools?" Urban Institute, November 28, 2018, https://www.urban.org/features/segregated-neighborhoods-segregated-schools.

ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places.

500 L'Enfant Plaza SW
Washington, DC 20024

www.urban.org