

OCTOBER 2020

A New Method for Describing Chicago Neighborhoods



TABLE OF CONTENTS

- 1 Introduction
- 2 The Importance of Neighborhoods
in Prior Research
- 4 Methods
- 5 Results of our Analyses:
The Five Neighborhood Groupings
- 10 Conclusion
- 11 Endnotes
- 13 Appendices
- 18 Acknowledgements

Authors

Information in this brief is based on research by:

Maia C. Connors, Isabel C. Farrar, & Amanda G. Stein

Start Early

Stacy B. Ehrlich, John Francis, & Sarah E. Kabourek

NORC at the University of Chicago

John Q. Easton

The University of Chicago Consortium on School Research

Cite as: Easton, J.Q., Francis, J., Farrar, I.C., Stein, A.G., Connors, M.C., Ehrlich, S.B., & Kabourek, S.E. (2020). *A new method for describing Chicago neighborhoods*. Chicago, IL: Start Early and NORC at the University of Chicago.

Introduction

Neighborhoods have enduring impacts on the people who live in them. They affect a wide range of social and personal experiences of their residents – from schooling, to access to health care and quality food, to social relationships and supports, and to civic participation. Therefore, understanding the characteristics of neighborhoods is important for informing policy decisions about how to most equitably and efficiently allocate services, supports, and resources. But determining how to characterize neighborhoods is not a straightforward task; they are multi-faceted, complex units made up of disparate factors such as people, physical environment, cultural norms, institutions, and businesses, just to name a few.

The process and results of defining neighborhoods may look different depending on the purpose. In this brief, we focus on how we – a team of education researchers – defined neighborhoods in Chicago. The initial goal of this work was to enable us to answer questions about how access to and enrollment in school-based pre-kindergarten (pre-k) may have varied by neighborhood characteristics. In conducting this work, we found the question of how to describe neighborhoods particularly interesting.

We used a data-driven method for characterizing neighborhoods that leveraged publicly available census data and allowed us to consider many neighborhood characteristics simultaneously. This method resulted in a parsimonious set of five neighborhood groupings in Chicago that enabled us to simplify how we understood the relationship of neighborhood char-

acteristics to a host of educational and other outcomes. In fact, our school district colleagues recommended that we produce this brief as a resource to those seeking to apply similar approaches in their work.

This brief shares our approach to defining neighborhood groupings and their characteristics and our findings about them. First, we review prior neighborhood research and ways of characterizing Chicago community areas and neighborhoods, outlining how the current work builds on these previous efforts. We then present our data-driven method for grouping neighborhoods and describe our neighborhood groupings. Finally, we discuss the potential uses and implications of our work. We hope that others in Chicago (or elsewhere) find this approach to defining neighborhood groupings useful – both for conducting research studies and providing services or resources.

The Importance of Neighborhoods in Prior Research

Many researchers from disparate fields such as education, sociology, and psychology have studied the relationship between neighborhood characteristics and children's outcomes. For example, much research has found a direct link between neighborhood poverty and poorer academic, behavioral, and health outcomes.¹ Other work has detailed the diverse mechanisms – including social relationships, norms, and institutional resources such as preschool quality – through which neighborhood characteristics can positively shape outcomes as well.² The resounding conclusion from these decades of research is that where we live, what is located near us, and who our neighbors are matter for many aspects of our lives.

Seminal Work on Defining Neighborhoods in Chicago

A classic study conducted by Robert Sampson (and many colleagues), called the Project on Human Development in Chicago Neighborhoods, identified many relationships between neighborhood characteristics and the individual and collective experiences and outcomes of their residents. The book that resulted from this ambitious study, *Great American City: Chicago and the Enduring Neighborhood Effect*³, provides scientific evidence of the impact of neighborhoods.

Sampson does not provide a single definition of the term “neighborhood,” but he does point out important themes that run through many historical definitions. One theme is location – neighborhoods are geographic units that are embedded in larger units such as cities. The ability for residents to interact in person may also be a feature of a neighborhood, meaning they are usually small enough units to allow residents to do so. A second is identity and connections – they are

characterized by a sense of social identity that is often defined by factors beyond location, such as race, ethnicity, and social class, or by the combinations of these factors. Neighborhoods may also gain an identity from what they are **not** – not poor, for example.

In Chicago, there are 77 official community areas, which have been in continual use since the early 1920's when they were first created.⁴ However, although these community areas connect geographically to specific locations in the city and have defined boundaries, they are geographically too big to be considered neighborhoods in Sampson's terms. In this sense, Chicago's 77 community areas do not define the city's neighborhoods because there is much diversity within many community areas and great variability from one neighborhood to another within a given community area. This suggests that to better understand their characteristics, neighborhoods in Chicago should be defined on a smaller scale.

In contrast, Sampson and his research colleagues identified 343 neighborhoods in Chicago by combining two to three adjacent census tracts (out of a total of 866) from the 2000 U.S. Census. Census tracts are small (typically they contain just 1,200 to 8,000 residents), relatively homogeneous geographic units used by the Census Bureau to present statistical data.⁵ These census tracts were not combined arbitrarily; the researchers considered boundaries (highways, rivers, railroad tracks) and demographic characteristics of the nearby census tracts to ensure that they combined relatively similar tracts. Sampson's approach helped to identify Chicago's individual neighborhoods with much greater specificity than is captured by the official community areas. Yet, this greater specificity and accuracy also created a new problem – 343 neighborhoods are harder to understand, analyze, and create policy for than are 77 community areas.

How the Current Work Builds On Prior Examinations of Neighborhoods

To account for this complication, our approach to grouping neighborhoods differs from Sampson's in a fundamental way. Rather than grouping census tracts together based on geographic proximity, we are defining "neighborhood groupings" based on shared sociodemographic characteristics – emphasizing Sampson's second definitional theme, identity, and connections over location. That is, our groupings include census tracts where people share many similarities in terms of race/ethnicity, income, and other characteristics, but they may have scattered locations. To do so, we used a data driven approach (**detailed in Appendix A**) to group the 797 census tracts in Chicago defined by the 2010 U.S. Census that had residents with similar characteristics.

We took an analytic approach that is referred to as a "person-centered" approach (or "neighborhood-centered" in this case) as opposed to the more typical "variable-centered" approach used by most prior studies of neighborhoods. In the latter case, neighborhood characteristics like those listed above are explored independently as predictors of relevant outcomes. For example, we might explore whether the concentration of Hispanic or Black residents in children's home neighborhoods is predictive of their enrollment in pre-k, holding constant levels of neighborhood poverty, language status, etc. While helpful in certain circumstances, trying to "disentangle" often-related neighborhood characteristics leaves us with results that are difficult to interpret. With our neighborhood-centered approach, we can make fewer and more intuitive comparisons among neighborhoods by considering a multitude of neighborhood characteristics simultaneously to understand how these variables, viewed in combination, relate to outcomes of interest. For example, using a neighborhood-centered analysis, we can compare pre-k enrollment patterns across a small number of different kinds of neighborhood groupings.

Although neighborhood-centered approaches are not very common in education research, we are not the first to attempt this work. For example, researchers have used this method for grouping neighborhoods to look at how neighborhoods are related to health outcomes.⁶ There is also a small number of examples from psychology.⁷ For example, one study used census data from 1990 to categorize neighborhoods based on several dimensions, including violence, disadvantage, and collective efficacy, to explore how neighborhood groupings were associated with adolescent antisocial behavior.⁸ Like ours, these studies all drew from large datasets (such as the census) to characterize neighborhoods based on factors other than location. Unlike our work, however, these studies tended to focus on older children and did not consider educational outcomes (with one exception⁹). Importantly, most prior work has grouped neighborhoods based on measures of structural or relational (dis)advantage (e.g., housing problems, green space, neighborhood disorder, violence, etc.). *Instead, we are seeking simply to describe neighborhoods in terms of the people who live in them.*

We conducted this investigation of how to categorize neighborhoods into meaningful groupings in order to facilitate a research study that examined the relationship between access to Chicago Public Schools pre-k classrooms and students' actual enrollment.¹⁰ Our research questions asked not only who enrolled in pre-k but also asked how geographic location and – more importantly – how the neighborhood context of children's residences was related to pre-k access and enrollment among different student groups. Thus, we were focused on several descriptors of residential neighborhoods that are important for the district to understand when making decisions about how best to support students' enrollment and success in school. These included typical indicators such as race/ethnicity and income and employment of residents,¹¹ as well as the prevalence of bilingual speakers and other linguistic characteristics of residents in neighborhoods. In addition, given previous research

in Chicago^A we considered the education level and occupation of residents^B within neighborhoods. This enabled us to examine access and enrollment patterns within and across different groupings of neighborhoods comprised of people like each other on these characteristics. Understanding neighborhoods based on individual resident characteristics (as opposed to structural or environmental characteristics) is useful in

two ways. First, it helps to characterize the contexts in which students live and thus the ways that their residential neighborhood might impact their outcomes and experiences in school (as we describe above). And second, it can serve as a “shorthand” method for identifying locations in which one is likely to find high concentrations of students who would benefit from similar kinds of supports and services.

Method

The Census Variables Used to Group Census Tracts Into Neighborhood Groupings

We used 12 different variables at the tract level from the from the 2012 American Community Survey 5-year estimates to conduct the analysis^C: four measure race/ethnicity; four measure language and place of birth; two measure income level and employment

(combined into one variable); and two measure education and occupation (combined into one variable). The technical names and table location within the American Community Survey files of each variable are contained in **Table B.1 in Appendix B**. The actual analysis used only ten variables, as the income and employment variables were combined into one, as were the education and occupation variables.

RACE/ETHNICITY

- Percent Asian
- Percent Black
- Percent Hispanic (non-White)
- Percent White

LANGUAGE AND PLACE OF BIRTH

- Percent Who Speak English Well
- Percent Bilingual
- Percent Who Speak Only Another Language (not English)
- Percent Foreign-Born

INCOME AND EMPLOYMENT

(Combined Into One Variable^D)

- Percent of Families with Income Above the Poverty Level
- Percent of Employed Males

EDUCATION AND OCCUPATION

(Combined Into One Variable)

- Mean Level of Education in Years (over 25 years old)
- Percent Employed as Management, Professionals

A We used two variables created at the University of Chicago Consortium on School Research: Income and Employment (combined into one variable) and Education and Occupation (combined into one variable). Each of these is composed of the two census items noted in the text. Documentation for the variables can be found in **Bryk, Sebring, Allensworth, Luppescu, & Easton (2010)**. They have been used regularly and successfully since then.

B **Brooks-Gunn, J., Duncan, G.J., Leventhal, T., & Aber, J.L. (1997).** *Lessons learned and future directions for research on the neighborhoods in which children live.* **Brooks-Gunn, J., Duncan, G.J., and Aber, J.L. (eds.)** *Neighborhood Poverty*, 1, 279-297.

C The American Community Survey is a product of the **US Census Bureau** that provides vital information on a yearly basis about the nation and its people through a nationally, state representative survey. Neighborhood type results from the 2012 five-year estimates were verified, and confirmed, by re-running analyses with the 2015 five-year estimates. 2012 was chosen because it was the year that made the most sense for the time period of interest of the Pre-K study.

D The income and employment variables are combined and negatively coded to create one variable.

Why We Chose These Specific Census Variables

We chose variables most aligned with the intent of expanding pre-k access for students most likely to benefit from pre-k enrollment. These included “high priority” students – students of color, those speaking a language other than English, and those living in neighborhoods with lower income and higher unemployment. Several policies prioritized these students and the neighborhoods in which they lived; we therefore wanted to identify neighborhoods by using similar indicators that the policies attended to. We also included a measure of education level and employment that has proven to be related to positive student outcomes in prior UChicago Consortium research.¹²

Analytic Technique Used to Group Census Tracts

Our groupings of census tracts were defined by using an advanced statistical technique^E that identifies neighborhoods that are similar in terms of the 10 census variables listed above and captures complexities of neighborhoods better than other statistical approaches. This is desirable from a research and analytic perspective given that it results in fewer variables to consider in our analyses. These new groupings are also easier to understand and describe. The statistical technique is a special case of a family of methods called “mixture models” (see **Appendix A** for statistical detail and software code).

There are two main goals of this type of analysis: **(1)** Create groups in which the *within-group similarities* in census characteristics are maximized, while at the same time, **(2)** Maximize the *between-group differences*. That is, each grouping of census tracts will contain the other tracts that are most like them and exclude the ones least like them.

Results

Our work resulted in the identification of *five distinct groupings* of census tracts in Chicago, some of which are scattered across the city with others more geographically concentrated. The five groupings are relatively easy to describe, and they are easily understood by those familiar with Chicago. We found that using the five neighborhood groupings added considerable value to our research project and made our findings easier to interpret than if we had used the “variable-centered” approach.

Note that Chicago’s longstanding residential segregation patterns are reflected in these neighborhood groupings. For a comprehensive examination, we highly recommend the following report from the Institute for Research on Race and Public Policy at the

University of Illinois Chicago: *A Tale of Three Cities: The State of Racial Justice Report in Chicago*.¹³ Here is an extensive quotation from the summary of that report:

“The central finding of this report is that racial and economic inequities in Chicago remain pervasive, persistent, and consequential. These inequities affect the lives of Chicagoans in every neighborhood; they have not just spatial but also deep historical roots and are embedded in our social, economic, political institutions; and they have powerful effects on the experiences and opportunities of all Chicagoans. The patterns are stark, if not entirely surprising. Chicagoans of all racial and ethnic groups want to live in safe and healthy communities where they don’t just subsist or survive, but not all have equal access.”

^E The technical name for this analytic method is latent profile analysis; see Appendix A for details.

That report provides a deep look into the context of Chicago neighborhoods and it also reflects one of our motivations for undertaking this effort to describe those neighborhoods in terms of who the residents are. We acknowledge that our groupings, as described below, reflect these long-standing inequities. Highlighting the sociodemographic characteristics of the residents in these neighborhoods is intended to provide useful information that can inform policy and supports to all communities across the city.

Five Chicago Neighborhood Groupings Descriptions

The five neighborhood groupings are as follows (listed in order of highest percentage of census tracts included in each). We deliberately chose not to give them descriptive names, given that the intent of this brief is to highlight the technique and purpose for grouping similar neighborhoods in Chicago, rather than naming them. We leave it to readers to choose the most appropriate names for their specific purposes.

■ **Group 1.** (31% of census tracts^F). This neighborhood group contains tracts almost entirely comprised of Black residents; nearly all residents are native-born and speak English well. These neighborhoods have the lowest proportion of households with incomes above the federal poverty level, employed males, and individuals in managerial jobs.

■ **Group 2.** (23% of census tracts). This neighborhood group contains tracts with high percentages of White residents, relatively few Black and Hispanic residents, and some Asian residents. Most residents in these tracts speak English well, few are bilingual or other language speakers, and few are foreign-born. These neighborhoods have the highest proportion of households with incomes above the federal poverty level, employed males, individuals in managerial jobs, and highest average years of education.

■ **Group 3.** (19% of census tracts). This neighborhood group contains tracts with high percentages of Hispanic residents and relatively few Black, White, and Asian residents. About two-thirds of residents in these tracts speak English well (a small proportion relative to the city of Chicago), and more than one third of residents are bilingual, speak a language other than English, and/or are foreign-born. In these neighborhoods, the proportion of households with incomes above the federal poverty level is similar to Chicago's city-wide average. These tracts have higher than average male employment rates, but residents are less likely to hold jobs in managerial roles, and have the lowest average years of education.

■ **Group 4.** (18% of census tracts). This neighborhood group contains tracts in which almost half of residents are White, and another third are Hispanic. These tracts also have the largest proportion of Asian residents relative to tracts in other neighborhood groups and Chicago's city-wide average. About three-quarters of residents in these tracts (a small proportion relative to the city of Chicago) speak English well, and many are bilingual, speak a language other than English, and are foreign-born. In these neighborhoods, the proportions of households with incomes above the federal poverty level and employed males are higher than Chicago's city-wide average. These tracts are similar to the city-wide average in terms of years of education and percentage of individuals in managerial jobs.

■ **Group 5.** (9% of census tracts). This neighborhood group contains tracts that are most racially diverse; they are half Black, one-fifth White, and one-fifth Hispanic, on average. Tracts in this group are similar to the Chicago-wide average on all variables, although slightly fewer residents than average are bilingual or speak a language other than English. These neighborhoods have a slightly smaller proportion of households with incomes above the federal poverty level, and more residents employed in managerial jobs.

^F This is percent of census tracts, not percent of the population. This is because the size of census tracts vary greatly.

Details About the Five Neighborhood Groupings

Table 1 provides a breakdown of census variables by neighborhood grouping and the average for census tracts within Chicago.

TABLE 1
Average neighborhood group population characteristics (c. 2012).

Variable	Group 1 (31%)	Group 2 (23%)	Group 3 (19%)	Group 4 (18%)	Group 5 (9%)	City of Chicago Census Tract Average %
Black	94.4%	5.5%	5.5%	6.7%	52.0%	37.8%
White	2.1%	75.3%	13.5%	47.4%	20.4%	30.8%
Hispanic	2.6%	10.6%	78.5%	30.8%	20.7%	25.3%
Asian	0.4%	6.9%	2.4%	13.4%	5.8%	5.05%
Speak English Well	98.6%	93.7%	60.1%	76.5%	88.8%	85.4%
Bilingual	2.9%	14.1%	37.5%	31.6%	16.9%	19.1%
Speak Only Another Language (not English)	1.7%	6.4%	39.5%	19.3%	10.9%	14.4%
Foreign-Born	2.2%	14.1%	40.2%	32.7%	17.0%	18.8%
Income (Families with Income Above the Poverty Level)	69.3%	95.0%	77.7%	86.3%	76.0%	80.4%
Employment (Employed Males)	53.2%	88.2%	79.2%	81.1%	71.9%	72.8%
Education (Average Level of Education in Years)	12.8	15.6	10.9	13.3	13.5	13.2
Occupation (Employed as Management)	25.6%	57.8%	15.9%	35.3%	37.2%	34.0%

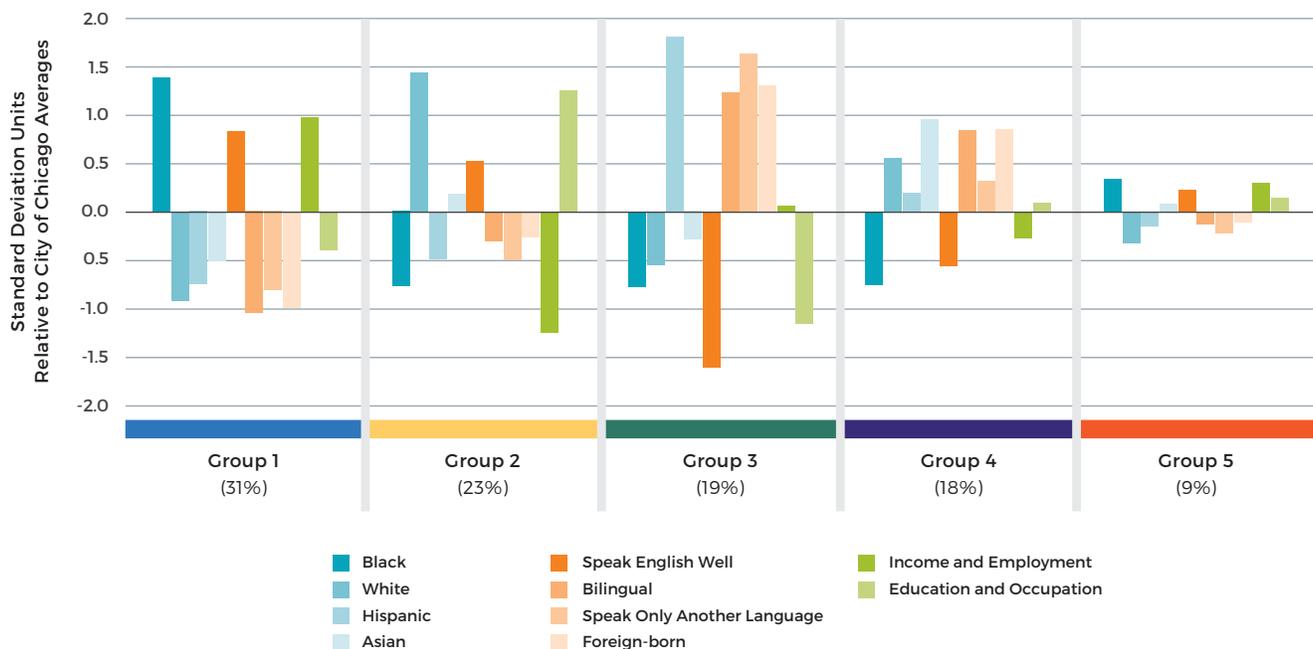
Figure 1 below provides values – in standard deviation units – for the 10 different census variables across the five neighborhood groupings (recall that Income and Employment were combined into one variable as were Education and Occupation). In this figure, each neighborhood grouping has ten bars – one for each of the 10 census variables used in the analysis. Bars above the zero line have higher values than the Chicago average for that variable, and correspondingly, bars below the zero line have lower values than the Chicago average for that variable. The graph uses “standard deviation units” where zero is equal to the average value, and +1.00 SD units is roughly equivalent to the 84th percentile ranking, and -1.00 SD units is roughly equivalent to the 16th percentile ranking.

To provide an example for understanding Figure 1, we “talk through” one neighborhood grouping – Group 3. This is the third most prevalent neighborhood grouping in Chicago and comprises 19% of census

tracts. It is also the neighborhood grouping with the most extreme values on six of the ten variables. These six (the highest and lowest bars) are Percent Hispanic, (1.81 SD above the city average), Speak Other Language (1.63 SD above average), Speak English Well (1.62 SD below average), Foreign-Born (1.31 SD above average), Bilingual (1.22 SD above average), and Education and Occupation (1.17 SD below average). The three other race/ethnicities – Black, White, and Asian – are also below the city average (at 0.79 SD, 0.56 SD, and 0.30 SD below, respectively). Only one combined variable, Income and Employment, is almost the same as the city average with a value of 0.05 SD units. **See Table D.1 in Appendix D** for precise standard deviation values for all neighborhood groupings.

The information contained within Table 1 and Figure 1 informed the narrative descriptions of the five neighborhood groupings above.

FIGURE 1
Characteristics of the five neighborhood groupings relative to Chicago averages.



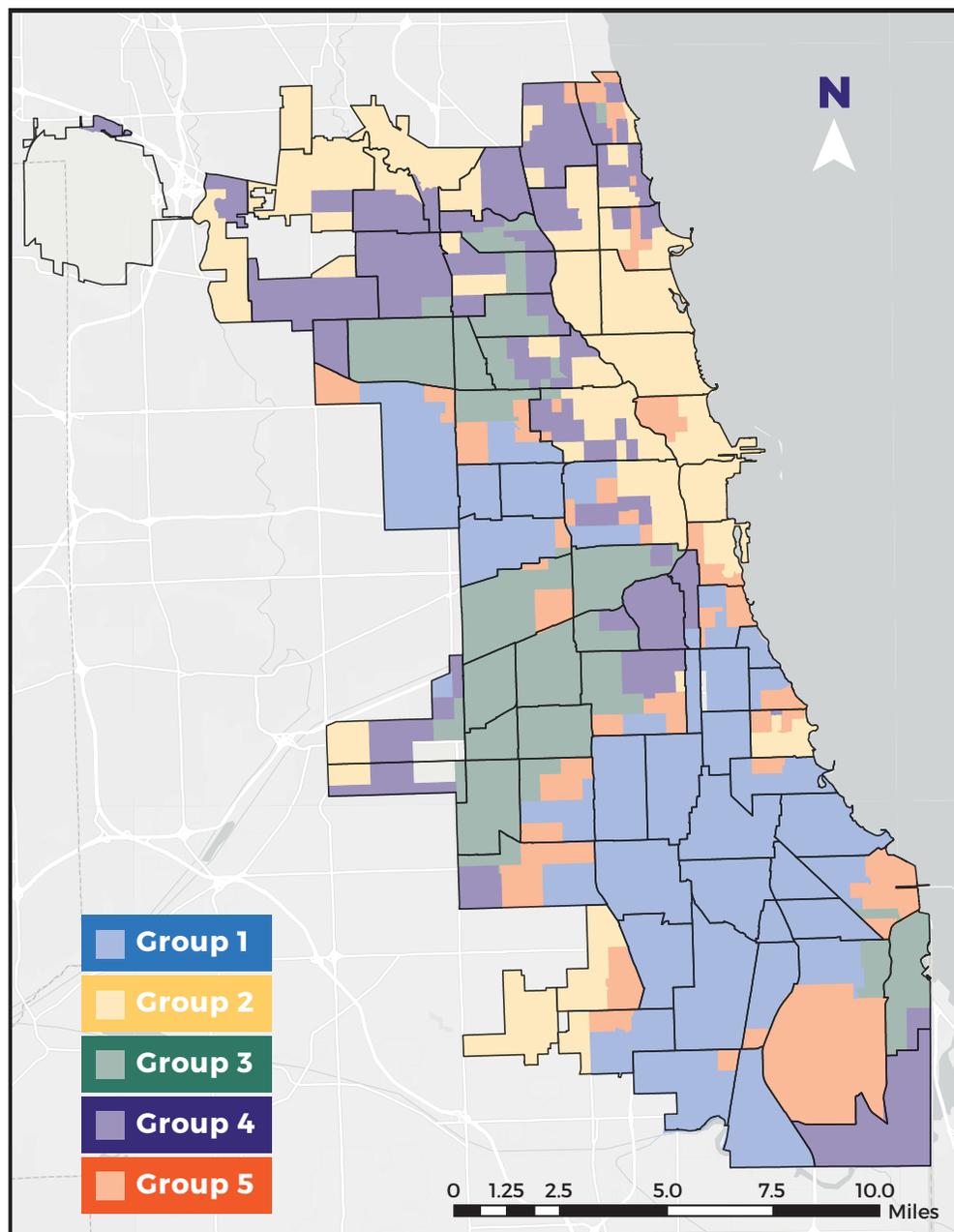
Our Neighborhood Groupings Mapped Onto the 77 Chicago Community Areas

The geographic locations of the five neighborhood groupings are shown in the Chicago map below, which also shows the boundaries of its 77 community areas. Those who know Chicago will recognize in **Figure 2**

that large swaths of the south and west sides are Group 1 neighborhoods; that many north-side, lake-front neighborhoods are Group 2 neighborhoods; and that Group 3 neighborhoods are found near the north-west and southwest areas of the city. Also, there are many Group 4 and Group 5 neighborhoods spread across the city, especially in between more homogenous areas of the city.

FIGURE 2

Map of Chicago with the five neighborhood groupings mapped onto the 77 community areas.



Conclusion

This “neighborhood-centered” rather than a “variable-centered” approach to the analysis enabled us to understand how neighborhoods in Chicago can be thought of in terms of the people who live in them. Our work provides just one example of how these methods can be useful to cities, school districts, and policy makers across the country. Future work can apply a similar approach to meet different aims.

Using this method for creating neighborhood groupings has aided our own research study in several ways. First, we can look at the city in a more fine-grained way by combining similar census tracts rather than the more typical use of Chicago’s 77 community areas, allowing us to see variation within community area more clearly. We can also examine outcomes (e.g., pre-k access and enrollment) more easily by neighborhood groupings than we could have with a more conventional approach that would have been more complicated and difficult to interpret. This is because we are less interested in how the 12 census variables independently relate to child outcomes and are more interested in how the combined variables defined neighborhood groupings and how the neighborhood groupings influence outcomes.

But ours is just one application of this approach. The aims of future work using similar methods will differ. For example, in this analysis we used data from the 2012 American Community Survey and replicated the analysis using 2015 data and found nearly identical

results. This analysis could easily be re-run with the most recent census data from the 2018 American Community Survey, to track how neighborhoods shift over time. It is also important to note that a similar analysis could be conducted using different census variables or by using other geo-coded data from different sources. Because we aimed to support the school district in understanding and improving students’ access and enrollment in pre-k, our work focused on describing neighborhoods in terms of the people who live in them. However, one could easily envision creating neighborhood groupings for other purposes that instead describe the neighborhood’s physical environment or resources, such as the presence of park land, playgrounds, libraries, museums, and other cultural institutions. Work focused on public health might include the presence of grocery stores and community gardens, or air quality indicators.

The opportunities for describing neighborhoods are great and should be tailored to the specific decisions to be made and research applications of users.

Endnotes

- 1 Leventhal, T., & Brooks-Gunn, J. (2000).** The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin*, *126*(2), 309.
- Katz, L.F., Kling, J.R., & Liebman, J.B. (2001).** Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, *116*(2), 607-654.
- Leventhal, T. & Brooks-Gunn, J. (2003).** Moving to opportunity: an experimental study of neighborhood effects on mental health. *American Journal of Public Health*, *93*(9), 1576-1582.
- Rosenbaum, J.E., Reynolds, L., & DeLuca, S. (2002).** How do places matter? The geography of opportunity, self-efficacy and a look inside the black box of residential mobility. *Housing Studies*, *17*(1), 71-82.
- McWayne, C.M., McDermott, P.A., Fantuzzo, J.W. & Culhane, D.P. (2007).** Employing community data to investigate social and structural dimensions of urban neighborhoods: An early childhood education example. *American Journal of Community Psychology*, *39*(1-2), 47-60.
- Thompson, F.T. (2002).** Student achievement, selected environmental characteristics, and neighborhood type. *The Urban Review*, *34*(3), 277-292.
- Plybon, L.E. & Kliewer, W. (2001).** Neighborhood types and externalizing behavior in urban school-age children: Tests of direct, mediated, and moderated effects. *Journal of Child and Family Studies*, *10*(4), 419-437.
- Jones, M. & Huh, J. (2014).** Toward a multidimensional understanding of residential neighborhood: A latent profile analysis of Los Angeles neighborhoods and longitudinal adult excess weight. *Health & Place*, *27*, 134-141.
- Humphrey, J.L., Lindstrom, M., Barton, K.E., Shrestha, P.M., Carlton, E.J., Adgate, J.L. & Root, E.D. (2019).** Social and Environmental Neighborhood Typologies and Lung Function in a Low-Income, Urban Population. *International Journal of Environmental Research and Public Health*, *16*(7), 1133.
- 2 Leventhal & Brooks-Gunn (2000).**
 - Dupéré, V., Leventhal, T., Crosnoe, R., & Dion, E. (2010).** Understanding the positive role of neighborhood socio-economic advantage in achievement: The contribution of the home, child care, and school environments. *Developmental Psychology*, *46*(5), 1227.
 - McCoy, D.C., Connors, M.C., Morris, P.A., Yoshikawa, H., & Friedman-Krauss, A.H. (2015).** Neighborhood economic disadvantage and children's cognitive and social-emotional development: Exploring Head Start classroom quality as a mediating mechanism. *Early Childhood Research Quarterly*, *32*, 150-159.
- 3 Sampson, R. J. (2012).** *Great American city: Chicago and the enduring neighborhood effect*. University of Chicago Press.
- 4 Zangs, Mary (2014).** *The Chicago 77: A Community Area Handbook*. Arcardia Publishing.
- 5 U.S. Census Bureau.** Census Glossary. Retrieved from: https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13; Chicago tracts range from 0 to 10,000.
- 6 Jones & Huh (2014); Humphrey et al. (2019).**
- 7 Abner, K.S. (2014).** Dimensions of structural disadvantage: A latent class analysis of a neighborhood measure in Child Welfare data. *Journal of Social Service Research*, *40*(1), 121-134.
Anderson, S., Donlan, A.E., McDermott, E.R. & Zaff, J.F. (2015). Ecology matters: Neighborhood differences in the protective role of self-control and social support for adolescent antisocial behavior. *American Journal of Orthopsychiatry*, *85*(6), 536.
Coley, R.L., Votruba-Drzal, E., Collins, M.A., & Miller, P. (2014). Selection into early education and care settings: Differences by developmental period. *Early Childhood Research Quarterly*, *29*(3), 319-332.
- 8 Anderson et al. (2015).**
- 9 Coley et al. (2014).**
- 10 Ehrlich, S.B., Connors, M.C., Stein, A.G., Francis, J., Easton, J.Q., Kabourek, S., & Farrar, I.C. (2020).** *Closer to Home: More Equitable Pre-K Access and Enrollment in Chicago*. Chicago, IL: University of Chicago Consortium on School Research.
- 11** On differential access to enrollment, **see Burgess, K., Chien, N., Morrissey, T., & Swenson, K. (2014).** *Trends in the use of early care and education, 1995-2011*. US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.
Dupéré et al. (2010).
Hastings, J.S., Van Weelden, R., & Weinstein, J. (2007). *Preferences, information, and parental choice behavior in public school choice* (No. w12995). National Bureau of Economic Research.
McCoy et al. (2015).
Nores, M., & Barnett, W.S. (2014). *Access to high quality early care and education: Readiness and opportunity gaps in America*. National Institute for Early Education and Center on Enhancing Early Learning Policy report. New Brunswick, NJ: Center on Enhancing Early Learning Outcomes. Accessed February, 13, 2015.

West, A. (2006). School choice, equity and social justice: The case for more control. *British Journal of Educational Studies*, 54(1), 15-33.

On differences in outcomes/effect of pre-k on high-priority student groups, see **Weiland, C., & Yoshikawa, H. (2013).** Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84(6), 2112-2130.

12 Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., & Easton, J.Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. University of Chicago Press.

13 Henricks, K., Lewis, A.E., Arenas, I., & Lewis, D.C. (2017). *A Tale of Three Cities: The State of Racial Justice in Chicago*. University of Illinois at Chicago Institute for Research on Race and Public Policy.

Appendix A

Details of our analytic method: Latent Profile Analysis Modeling

Modeling

To conduct the Latent Profile Analysis (LPA), our pre-k access and enrollment study team utilized MPLUS^G to run a mixture model with a maximum likelihood estimator.^H LPA is a special case of mixture modeling in which the latent profiles (or what we call “neighborhood groupings” above) explain the relationships among the observed continuous dependent variables (neighborhood variables about characteristics of residents) through a set of linear regression equations. The key modelling choices one has with an LPA are which variables to include, and how many different profiles the data should be grouped into. We ran models investigating what the data looked like with 2-8 profiles, upping the number of random starts as necessary to ensure the best loglikelihood replication and that we were not achieving a local maximum.^I Additionally, we ran two sets of models; one using variables pulled directly from ACS including a variable for poverty level (% of those living below 150% of the federal poverty level), and another using versions of the ACS variables standardized to the census tracts in the City of Chicago with two variables often calculated by the UChicago Consortium (Income and Employment combined into one variable and Education and Occupation Combined into one variable) in place of the poverty level variable.

Fit. In LPA, and structural equation models (SEMs) more broadly, there is quite a bit of subjectivity in choosing the model with the best fit. Below are the different indicators of model fit we used to help decide on which model (how many profiles) to move forward with. While some of the indicators provided justification for selecting 6 or 7 profiles, we ended up deciding on 5 profiles as this most closely aligned with our knowledge of Chicago, performed well on the fit indicators, all while maintaining a substantive percentage of tracts per group. From the classification probabilities shown in **Table A.1**, we can see that the 5-profile model does a particularly good job of defining profiles that are very different from one another (top to bottom diagonal).

- **Log-likelihood (LL) value:** Higher values (closer to 0) indicate better fit
- **Lo-Mendell-Rubin (LMR) likelihood ratio test:** Used to compare models with different numbers of clusters; significant p-value indicates that the more complex model (with more clusters) fits ‘better’
- **Adjusted BIC:** Smaller values indicate better fit; can compare non-nested models, but gives no p-value
- **Entropy:** Used to represent how well the posterior probabilities were collectively able to confidently classify individuals; Higher entropy indicates greater confidence

G The MPLUS code use for this analysis can be found in Appendix A.

H “Mixture modeling refers to modeling with categorical latent variables that represent subpopulations where population membership is not known but is inferred from the data. With continuous latent class indicators, the means of the latent class indicators vary across the classes as the default.” MPLUS User Guide Chapter 7

I In MPLUS for low numbers of classes we found that Starts= 2000 50 (representing the number of initial stage starts and number of final stage optimizations) was sufficient to ensure a stable solution, however it was necessary to ramp up the number of random starts to Starts= 2000 500 to ensure convergence for models with the highest number of classes. The creators of MPLUS recommend doubling the number of starts even after convergence as a check to make sure a local maxima was not reached, which we did

TABLE A.1**LPA model fit statistics and classification probabilities.**

Cluster	LL Value	LMR	Adj BIC	Entropy	% Sample per Class
2	-9125.136	4806.328, p<.01	18358.937	0.972	71, 29
3	-6937.886	4315.774, p<.01	14022.996	0.985	36, 37, 27
4	-6282.513	1293.149, p<.01	12750.808	0.975	25, 19, 21, 36
5	-5812.806	926.803, p<.01	11849.953	0.982	9, 31, 23, 19, 18
6	-5424.681	765.830, p<.1	11112.26	0.985	9, 22, 17, 31, 19, 2
7	-5071.515	696.849, p<.542	10444.487	0.974	15, 8, 20, 3, 11, 31, 12
8	-4806.424	523.064, p<.03	9952.864	0.981	4, 20, 32, 11, 14, 12, 6, 3

TABLE A.2**Classification probabilities for five cluster LPA model.**

	Most Likely Latent Class Membership				
	1	2	3	4	5
Final Latent Class	0.993	0	0.007	0	0
1	0	0.98	0.003	0.004	0.012
2	0.009	0.002	0.977	0.012	0
3	0	0.002	0.013	0.984	0
4	0	0.003	0	0	0.997
5					

2012 v 2015 Census Data

The same five neighborhood groupings were generated when using the 2012 vs. 2015 census data. However, roughly 12% of census tracts changed their group membership between 2012 and 2015. For this study on pre-k access and enrollment, it was determined to simply use the 2012 group classification, with robust-

ness checks being made using the 2015 data. Even though about 12% of census tracts changed classification, the percentage of tracts in each neighborhood type remained relatively stable, as can be seen in **Table A.2**.

TABLE A.2
Census tract distributions using 2012 vs 2015 data.

2012 Neighborhood Groupings	2015 Neighborhood Groupings					Total
	Group 1	Group 2	Group 3	Group 4	Group 5	
Group 1	234	2	2	0	13	251
Group 2	3	170	0	9	1	183
Group 3	1	0	137	9	2	149
Group 4	2	9	13	117	2	143
Group 5	7	6	4	7	47	71
Total	247	187	156	142	65	797

MPLUS CODE (FOR ABOVE ANALYSIS)

TITLE: Census latent profile analysis;
DATA: FILE IS data/2015lpa_zscored.csv;
VARIABLE: NAMES ARE Id2 pblack pwhite phisp
 pasian pengwell pbiling pothlang
 pfborn incemp edocc;
 USEVAR ARE pblack pwhite phisp pasian
 pengwell pbiling pothlang pfborn incemp
 edocc;
 MISSING ARE ALL (9999);
 CLASSES = c(5);
ANALYSIS: TYPE = MIXTURE;
 ESTIMATOR = MLR;
 Starts= 600 150;
SAVEDATA: FILE IS 2015/pprob_c5_zscored.dat;
 SAVE = CPROBABILITIES;

Appendix B

Variables

TABLE B.1

Variables used to create the five neighborhood groupings and corresponding ACS data file and definitions.

Variables Used for Five Neighborhood Groupings	ACS Data File (2012 5-year estimates)	Specific ACS Variable Definition or Calculation
Black	<p>Table S0601</p> <p>Selected characteristics of the total and native populations in the United States</p>	% Black or African American, not Hispanic or Latino
White		% White alone, not Hispanic or Latino
Hispanic		% Hispanic (non-white)
Asian		% Asian
Speak English Well	<p>Table S1601</p> <p>Language spoken at home</p>	% Speak only English (Speak English Well)
Bilingual		% Speak a language other than English + Speak English only or speak English "very well" (Bilingual)
Speak Only Another Language (not English)		% Speak a language other than English + Speak English less than "very well" (Speak Other Language)
Foreign-Born	<p>Table B05002</p> <p>Place of birth by nativity and citizenship status</p>	% Foreign-Born
Income (Families with Income Above the Poverty Level)	<p>Table B17010</p> <p>Poverty status in the past 12 months of families by family type by presence of related children under 18 years by age of related children</p>	% of families with income in the past 12 months at or above poverty level
Employment (Employed Males)	<p>Table B23022</p> <p>Sex by work status in the past 12 months by usual hours worked per week in the past 12 months by weeks worked in the past 12 months for the population 16 to 64 years</p>	% of Males Worked in the past 12 months
Education (Level of Education in Years)	<p>Table B15002</p> <p>Sex by educational attainment for the population 25 years and over</p>	Mean level of Education (in years)
Occupation (Employed as Management)	<p>Table C24010</p> <p>Sex by occupation for the civilian employed population 16 years and over</p>	% employed as management, business, science, and arts occupations

Appendix C

Additional Details About Results

TABLE C.1

The precise values in standard deviation units used in Figure 1 in the main text.

Variable	Group 1 (31%)	Group 2 (23%)	Group 3 (19%)	Group 4 (18%)	Group 5 (9%)
■ Black	1.375	-0.785	-0.787	-0.758	0.346
■ White	-0.938	1.446	-0.565	0.541	-0.335
■ Hispanic	-0.772	-0.497	1.81	0.192	-0.162
■ Asian	-0.535	0.21	-0.301	0.945	0.081
■ Speak English Well	0.84	0.526	-1.621	-0.574	0.221
■ Bilingual	-1.066	-0.323	1.22	0.836	-0.153
■ Speak Only Another Language (not English)	-0.82	-0.516	1.633	0.315	-0.232
■ Foreign-Born	-1.017	-0.285	1.314	0.847	-0.122
■ Income and Employment	0.973	-1.257	0.054	-0.293	0.294
■ Education and Occupation	-0.319	1.258	-1.168	0.09	0.136

Acknowledgements

The authors would like to acknowledge the partnership, support, critical feedback, and encouragement from many of our colleagues. First, this research would not have been possible without the support of our dedicated colleagues, both former and current, at Chicago Public Schools (CPS; Michael Abello, Anna Colaner, Sarah Dickson, Leslie McKinley, Bryan Stokes, and Noriko Wodzien), the City of Chicago Office of the Mayor (Samantha Aigner-Treworgy, Jennifer Alexander, Cara Bader), and the Department of Family and Support Services (Elizabeth Stover).

In particular, Sarah Dickson (Director of External Research at CPS) suggested that we write this brief to describe how we created Chicago neighborhood groupings as part of a research study on the relationship between access and enrollment in school-based pre-kindergarten in Chicago. Furthermore, a team of Chicago early childhood education administrators, led by Bryan Stokes (Chief of Early Childhood Education at CPS) partnered with our research team to provide excellent feedback on framing and wording choices that were descriptive and asset-based when possible. Additionally, Caroline Ebanks (our program officer at the Institute of Education Sciences), encouraged us to develop this written product describing how we applied these innovative statistical methods in

our study. Our advisory board member, Dana Charles McCoy (Assistant Professor, Harvard Graduate School of Education) provided critical guidance on statistical and technical aspects of these analyses, as did members of the University of Chicago Consortium on School Research “data group.”

Finally, a huge thank you to Kim Ptak (Start Early), who oversaw the design, layout, and production of this brief, and to Jeff Hall for his incredible design and layout work on this publication. In addition, we appreciate the support of our marketing and communications colleagues Jessica Tansey (UChicago Consortium on School Research) and Eric Young (NORC) for their help with coordination on this publication and dissemination of these research findings.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through grant #R305A180510 to NORC at the University of Chicago. We are especially grateful to our program officer, Caroline Ebanks.

The opinions expressed in this document are those of the authors and do not necessarily represent views of Start Early, NORC at the University of Chicago, the University of Chicago Consortium on School Research, the Institute of Education Sciences, or the U.S. Department of Education.



33 West Monroe Street
Suite 1200
Chicago, IL 60603

312.922.3863
StartEarly.org

f /startearly

t @startearly

in /startearly

@ /startearly

ABOUT

Start Early (formerly known as the Ounce of Prevention) is a nonprofit public-private partnership advancing quality early learning and care for families with children, before birth through their earliest years, to help close the opportunity gap. For nearly 40 years, Start Early has delivered best-in-class doula, home visiting and Early Head Start and Head Start programs. Bringing expertise in program delivery, research and evaluation, professional development and policy and advocacy, Start Early works in partnership with communities and other experts to drive systemic change so that millions more children, families and educators can thrive. Learn more at www.StartEarly.org.

NORC at the University of Chicago is an objective, non-partisan research institution that delivers reliable data and rigorous analysis to guide critical programmatic, business, and policy decisions. Since 1941, NORC has conducted groundbreaking studies, created and applied innovative methods and tools, and advanced principles of scientific integrity and collaboration. Today, government, corporate, and nonprofit clients around the world partner with NORC to transform increasingly complex information into useful knowledge.