Center for American Progress

Measuring Child Care Supply Using the Enhanced Two-Stage Floating Catchment Area Method

By Rasheed Malik, Won F. Lee, Aaron Sojourner, and Elizabeth E. Davis

The distance-based approach to measuring child care supply and adjusting for nearby demand was introduced in a 2019 article published in the journal *Early Childhood Research Quarterly*.[1] Economists Davis, Lee, and Sojourner developed the methodology by adapting techniques from the field of geography in order to study the distribution of access to child care in Minnesota and adapted them for this project.[2] The advantage of this approach is that it alleviates key limitations of the area-based measure that has commonly been used in child care research literature. Area-based measures are limited due to the fact that the unit of analysis is based on administrative or political boundaries that are arbitrarily defined from the perspective of families needing child care for their children. Popular area-based measures of access to child care supply typically measure access using the ratio of the total capacity of licensed providers in an area divided by the estimated number of children in that area. Geographers have long noted concerns about the potential for information loss and statistical bias when using area-based measures.[3]

This new approach in the original journal article uses an enhanced two-stage floating catchment area method (E2SFCA). In the accompanying issue brief, this method is simply referred to as the distance-based approach to measuring child care supply. The distance-based approach differs from an area-based analysis in that it is centered on family locations and assumes families are interested in nearby providers whether or not they are located in the same census tract or other administrative area unit. The E2SFCA method measures the supply of slots within a family-specific catchment area and adjusts each nearby provider's capacity based on the number of young children near that provider who could potentially use the provider's slots. The two stages are, first, the catchment area around the provider's location and, second, the catchment area around the location of each family with young children.

In stage one of the two-stage calculation, a weighted capacity-to-population ratio is generated for every child care provider in the data set. For each provider, the model first identifies all family locations within the 20-minute drive-time catchment area around the provider's location. Drive times are computed taking speed limits and road networks into account so they tend to include more miles of distance in more-rural areas. The sum of children in these nearby families serves as a proxy for the amount of potential demand for the provider's capacity. However, since children closer to a provider are more likely to use that provider, children that are farther away are discounted using a Gaussian distance decay function.[4] The capacity (number of slots) of each provider is divided by the sum of the distance-discounted total number of children to obtain a capacity-to-nearby child population ratio for every provider. Ultimately, two providers with the same capacity will have different capacity-to-population ratios if one has more young children close by. A provider's slots will be less accessible to a given family if more children from other families live close to the provider and thus may compete for a spot.

Stage two of the methodology determines the quantity of local child care supply for each family location, based on the total capacity of nearby providers adjusted for their nearby young-child population. This is accomplished by identifying all providers within the 20-minute catchment area around the family's location and computing each one's slots-to-population ratio, discounting for travel time using the same distance-decay function as was used for the provider's capacity-to-population ratio. Summing these across the providers near each family yields the family's measure of access to nearby child care supply. This measure of supply, which measures supply nearby a family location adjusted for nearby population, increases if the family has more slots nearer by and decreases if more young children are nearer by those slots. This distance-based measure is similar to the area-based measure. However, it is centered on family locations and incorporates driving time as a measure of accessibility.

To construct the distance-based measure of families' access to child care, we need information on the residential locations of families with young children. We approximate family locations based on census block-group level child density information from the 2013–2017 five-year estimates from the American Community Survey (ACS). In operationalizing the E2SFCA method to compute nation-wide access measures, we first divided the nation into a grid of hexagons and used the hexagon centroids as a proxy for the family's location in the computation process. For ease of computation, we used one-quarter mile side hexagon centroids for metro areas, and 1-mile side hexagon centroids for nonmetro areas. By transforming the block-group level child density information from the ACS data into hexagons, we approximate the spatial distribution of families with young children across the nation.

Travel time between each hexagon centroid and each child care provider in the dataset is estimated using osrmtime, a Stata command that calculates driving time between two geocoded points.[5] For each hexagon centroid, providers within 20-minutes driving time are identified. Drive times are computed taking speed limits and road networks into account so they tend to include more miles of distance in more-rural areas.

Finally, for visualization, we generate a 10 percent population sample of synthetic family points to approximate the spatial distribution of families' residential locations. The likelihood of a synthetic family location being in a particular block group is proportional to the likelihood that a real family with children under age 6 lives in that block group. The exact location of any synthetic family within a block group is random assuming a uniform distribution of families within the block group boundaries. Impossible residential locations are restricted for synthetic families to exclude lakes, rivers, and state parks. These locations are generated in ArcGIS using the Geospatial Modelling Environment add-on tool, ultimately producing a geocoded dataset of synthetic family location points.[6]

While this study provides new distance-based measures of access to the supply of licensed child care, there are a number of limitations and directions for future research. An important limitation is that this research does not use data on actual residential locations of families who are seeking or using child care; therefore, there is potential for error in the locations and travel times. In addition, this analysis assumes travel from home to child care by car and is not able to incorporate travel times between work location and child care or by other modes of transportation. Despite these limitations, these new adjusted supply measures provide new insights into the supply and demand of early care and learning opportunities for families across the country.

Rasheed A. Malik is a senior policy analyst for Early Childhood Policy at the Center for American Progress. Won Lee is a research analyst at Stanford University. Aaron Sojourner is an associate professor at the University of Minnesota's Carlson School of Management. Elizabeth E. Davis is a professor of applied economics at the University of Minnesota.

Endnotes

1 Elizabeth Davis, Won Fy Lee, and Aaron Sojourner, "Family-centered measures of access to early care and education," Early Childhood Research Quarterly 47 (2) (2019): 472–486, available at https://www.sciencedirect.com/science/article/pii/S0885200618300851?via=ihub.

2 Elizabeth Davis, Won Fy Lee, and Aaron Sojourner, "Mapping Access to Child Care for Minnesota Families," available at http://childcareaccess.org (last accessed June 2020).

3 Shawna J. Dark and Danielle Bram, "The modifiable areal unit problem (MAUP) in physical geography," Progress in Physical Geography: Earth and Environment 31 (5) (2007), available at https://journals.sagepub.com/doi/abs/10.1177/0309133307083294?journalCode=ppga.

4 For more details on how this function operates, see the technical appendix in Davis, Lee, and Sojourner's "Family-centered measures of access to early care and education."

5 Stephan Huber and Christoph Rust, "Calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM)," The Stata Journal 16 (2) (2016): 416–423, available at https://journals.sagepub.com/doi/pdf/10.1177/1536867X1601600209.

6 Geospatial Modelling Environment, "Introducing the Geospatial Modelling Environment," available at http://www.spatialecology.com/gme/ (last accessed June 2020).