# TABLE OF CONTENTS

- Background and Overview ........................................................................................................... 1
- Data Sources ................................................................................................................................. 2
  - ECE program data ....................................................................................................................... 2
  - Early childhood population ........................................................................................................ 3
  - Residential location ................................................................................................................... 3
  - Family income ............................................................................................................................ 3
  - Roads and Public Transit Networks ............................................................................................ 3
- Data Preparation ............................................................................................................................ 3
  - ECE provider location and cost .................................................................................................. 3
  - Estimating population of children under age 6 at each residential lot ........................................ 4
  - Creating road map and transit network databases and catchment areas .................................... 4
- Creating Indexes .......................................................................................................................... 4
  - Spatial access and nearby seats index ....................................................................................... 4
  - Affordability index ...................................................................................................................... 5
  - Quality index .............................................................................................................................. 5
  - Aggregating indexes at the census tract level ............................................................................ 6
  - Combined indexes ...................................................................................................................... 6
  - Visualizing an overlay of three indexes ...................................................................................... 6
- Creating Maps ............................................................................................................................... 7
- References ...................................................................................................................................... 7
- Suggested Citation and Acknowledgments ..................................................................................... 8
BACKGROUND AND OVERVIEW

Access to affordable, conveniently located, and high-quality early care and education (ECE) is essential to the well-being of children, families, and communities. To measure accessibility, policymakers most often use metrics such as the number of licensed seats in a geographic region, average cost, or the percentage of centers at different levels of a quality rating system.

As a field, we need to conceptualize and measure access in a more nuanced way. A rich approach to measurement would be multidimensional, localized, and spatially family-centered (Friese et al., 2017; Lin & Madill, 2019).

- Multiple dimensions of access should be considered simultaneously. Families’ ECE choices are unlikely to be determined by a single consideration, e.g., open seats, cost, distance, or hours, but rather, by a combination of factors. Furthermore, social justice requires equity in all key dimensions of access.

- Access should be measured at the local level. Data at the state, county, or municipal levels are not sufficiently granular to reflect disparities across and within communities. Families live and conduct their daily business in a relatively small orbit, and constraints such as cost, distance, and open seats vary from neighborhood to neighborhood. Family characteristics, such as ethnicity, education, and income may also distinguish neighborhoods. From a policy perspective, the ability to identify “hot spots” of the highest need could result in a more strategic and effective allocation of resources. Finally, highly localized data would allow grassroots advocates to speak with additional authority about the needs of their communities.

- Data should be based on the distance between the home and ECE providers and not on static boundaries such as county lines, census tracts, or ZIP codes. Since families often cross these abstract boundaries to go to work or access services, metrics given in a radius around a family’s location offer a more authentic estimate of resource availability from a family’s perspective (Davis, Lee & Sojourner, 2019). The use of static boundaries is widespread in the ECE literature. In contrast, other fields, including public health and urban planning, use distance-based approaches to measure access to resources such as health care and parks (e.g., Li, 2016; Luo & Wang, 2003).

In this technical report, we present the methodology used in a study of equitable ECE access in the state of Hawai’i. Our purpose was to develop an innovative approach to measuring multiple components of ECE access within reasonable proximity of a family’s home. Our measures take into account commuting time or distance, as well as the number of children potentially competing for these nearby seats. As a result, the measures provide an accurate metric of the likelihood that a family can obtain ECE services in their community.

We developed the following indexes:

- **Spatial Access:** The demand-adjusted supply of ECE seats close to a child’s home, expressed as seats per child. This index represents the likelihood that a child can be seated. For this index, high scores are desirable.
• **Nearby Seats:** The demand-adjusted supply of ECE seats close to a child’s home expressed as children per seat, i.e., the inverse of spatial access. For this index, low scores are desirable.

• **Affordability:** The availability-weighted average cost of a nearby seat as a percentage of that area’s median family income. This index represents the typical cost burden relative to family income. For this index, low scores are desirable.

• **Quality:** The availability-weighted likelihood that a nearby seat is in a center with a national ECE accreditation or in a public preK classroom. For this index, high scores are desirable.

• **Combined Access:** The average of the standardized scores for spatial access, affordability, and quality, reflected as needed, so that high scores represent better access.

Indexes were calculated for three different catchment area scenarios: 5-mile driving distance, 10-mile driving distance, and a 30-minute public transit ride. These catchment areas were chosen based on stakeholder input on what constitutes a reasonable commute for urban and rural areas, and for families using private cars vs. public transportation.

Results were made available to community users via an interactive mapping website. When viewing each index, users can select among the three catchment areas (5- or 10-mile drive, 30-minute bus ride) and two levels of granularity (residential lot vs. census tract). We expect tract-level data to be more useful for most policy and planning purposes. The lot-level data should be useful for highly localized questions, such as specific locations within a census tract where new seats would serve the largest number of nearby children.

This project was intended to serve as a proof-of-concept and model for other states and municipalities. The methods are flexible and can be adjusted as needed to meet the needs of different users.

**DATA SOURCES**

**ECE PROGRAM DATA**

The ECE programs included in the project were registered family childcare homes, licensed group childcare homes, licensed infant-toddler and preschool centers, and public preK classrooms. The state Executive Office on Early Learning provided data on the location and capacity of public preK sites. The statewide Child Care Resource and Referral Agency (CCRRA) provided data on all other ECE programs. The CCRRA administrative data were current as of September 2019 and included the following fields: license number, license type, street address, licensed and desired capacity, ages served, tuition fees, and accreditation status. While the State of Hawai‘i Department of Human Services determines the license type and maximum site capacity, most other fields were based on provider self-report on an annual market rate survey.

---

1 Currently in Hawai‘i, buses are the only modes of public transportation.
2 Based on stakeholder input, we included the main providers of formal, regulated ECE services that offer a school-day or full-time schedule. ECE sectors not included were military childcare; early intervention; self-contained public special education preschool classrooms; family, friend, and neighbor care; and family-child interaction programs.
EARLY CHILDHOOD POPULATION
Population estimates for the number of children under age 6 were obtained from the American Community Survey (ACS) 2018 five-year estimates (U.S. Census Bureau 2019a). We excluded children who lived on military bases because data for on-base ECE facilities were not publicly available. There were an estimated 91,150 children age birth through five.

RESIDENTIAL LOCATION
Residential property information was extracted from a commercially available real property database that included the coordinates for each property, whether the property was a residential lot, and if so, the number of housing units on the lot (Digital Lightbox 2017). In total, there were 281,124 residential lots comprising over 458,000 housing units (e.g., single-family homes, discreet apartments inside a condominium or apartment building). These data were used to approximate the spatial distribution of young children’s residences.

FAMILY INCOME
Because no data source for individual family income exists, we used ACS five-year estimates of the median income of families with children under 18 as a proxy measure (U.S. Census Bureau, 2019b). We assumed that the family income at each residential lot was equal to that of the census tract in which the lot was located.

ROADS AND PUBLIC TRANSIT NETWORKS
Street centerline (road) data of each county were accessed through the state geospatial open database (Hawaii State Office of Planning, 2017). This database contained information on the direction of travel, speed categories, and access restrictions. We also obtained public transportation information from the Department of Transportation Services of each county. Transportation data were in the form of a GTFS feed that included transit agencies, routes, and schedules (i.e., stops, trips, and stop times). For an overview of GTFS feeds, see MobilityData (n.d.)

DATA PREPARATION
ECE PROVIDER LOCATION AND COST
Provider addresses were geocoded to the street level and assigned longitude and latitude coordinates using Google Geocoding API. To calculate each provider’s average fee for full-time monthly care, we first considered whether the provider charged a single or age-variable rate. In the latter case, tuition was the weighted average based on cost and enrollment for each age group served. If a provider reported age-based fees but not enrollment by age group, we assumed that the age distribution served was the same as that for other providers of the same type in the same county. Public preK and most Head Start and Early Head Start classrooms are free of charge; these programs were assigned average tuition of zero dollars. (Tuition was included for the small number of Head Start and Early Head Start classrooms that offered full-day care for a modest fee.) Note that tuition represented published market rates and did not

3 A total of 292 census tracts were used in to compute index scores. We omitted tracts that comprised military bases, uninhabited coastal areas with ACS population estimates of zero, and tracts with no residential properties.
reflect scholarships or needs-tested subsidies that could lower out-of-pocket costs for some families. In addition, not all families are eligible to enroll in Head Start, Early Head Start, or public preK.

ESTIMATING POPULATION OF CHILDREN UNDER AGE 6 AT EACH RESIDENTIAL LOT

To estimate the number of young children living at each residential lot, we used the ACS 5-year population estimates by census tract for children under age six (U.S. Census Bureau 2019a). We assumed that every residential lot housed young children, and that children were distributed to residential lots within a census tract in proportion to the number of housing units at each lot. The number of children under age six ($P_i$) per residential lot $i$ was approximated as follows:

$$ P_i = (H_i / H_b) \times P_b $$

where $P_b$ is the number of children under age 6 within census tract $b$, $H_i$ is the number of residential units at the lot, and $H_b$ is the number of residential units within census tract $b$.

CREATING ROAD MAP AND TRANSIT NETWORK DATABASES AND CATCHMENT AREAS

The first step in calculating catchment areas was to create a road and public transit-enabled network datasets, respectively. We created the road network dataset from the street centerline geodata using the ArcGIS Network Analyst extension tool. We created the public transit-enabled network dataset from the GTFS and street centerline data, again using the ArcGIS Network Analyst extension.

The second step was to use the road and public transit network datasets to create catchment area boundaries for each ECE provider, using the Generate Service Areas tool from the ArcGIS Pro Network Analysis toolset. We used the road network data to create driving distance catchment areas; these catchment areas contained all street surfaces falling within a given driving distance (5-miles and 10-miles, respectively) from the provider. The 30-minute public transit catchment included areas where a commuter could reach the provider within a 30-minute commute and arrive by 8 a.m.

CREATING INDEXES

SPATIAL ACCESS AND NEARBY SEATS INDEX

For each residential lot, the spatial access index was calculated as follows:

Step 1.1. For each provider $j$, we calculated the capacity-to-population ratio $R_j$ by summing the number of children age 6 at each lot within a defined catchment area $D$ from provider $j$:

$$ R_j = N_j = C_j / \sum_{i \in U_j} P_i $$

where $C_j$ is the capacity of provider $j$, $P_i$ is the number of children under age 6 at residential lot $i$, and $U_j = \{i : d(i, j) \leq D\}$ is a set of residential lots (i) within the defined catchment area $D$ (e.g., 5-mile, 10-mile, or 30-minute transit time) centered at a given provider $j$.

---

4 Because they are based on road layout, catchment areas are complex polygons, not symmetric circles centered on each provider.
Step 1.2. Some children will have access to more than one provider within the driving distance or public transit trip of interest. Thus, for each lot \( i \), the spatial access index \( A_i \) was calculated by summing \( R_j \) for all providers within the \( D \)-mile/time commute from a family’s location.

The equation is:

\[
A_i = \sum_{j \in U_i} R_j
\]  

(2)

where \( U_i = \{ j : d(i,j) \leq D \} \) is a set of providers \( j \) within the time or distance that defines catchment area \( D \).

The spatial access index \( A_i \) can be interpreted as the number of seats per child within the catchment area of interest. Because we were concerned that this ratio may not be immediately understood (i.e., that a spatial access score of .25 indicates there are enough seats to serve one out of four children), we also created a nearby seats index, which is simply the inverse of \( A_i \):

\[
N_i = 1/A_i
\]  

(3)

The nearby seats index can be interpreted as the number of children per seat within the catchment area of interest. For reader clarity, we used \( N_i \) on the mapping site. An important drawback of \( N_i \) is that it cannot be calculated with there are no nearby providers, i.e., when \( A_i = 0 \). For this reason, we used \( A_i \) to calculate combined index scores, and noted \( N_i = 0 \) as “Not applicable, no providers” on the website maps.

**AFFORDABILITY INDEX**

For each residential lot, the affordability index was calculated as follows:

Step 2.1. For each lot location \( i \), we calculated childcare cost \( S_i \) as the availability-weighted average tuition fees of nearby providers:

\[
S_i = \sum_{j \in U_i} (E_j * R_j / \sum_{j \in U_i} R_j)
\]  

(4)

where \( E_j \) is the monthly tuition of provider \( j \), \( R_j \) is the capacity-to-population ratio within the \( D \)-mile (or 30-minute) catchment area centered at provider \( j \), and \( U_i = \{ j : d(i,j) \leq D \} \) is a set of providers \( j \) that are within a \( D \)-mile (or 30-minute) commute.

Step 2.2. For each lot location \( i \), the affordability index \( F_i \) was computed as the ratio of the availability-weighted average tuition \( S_i \) to the annual median income of families at lot \( i \):

\[
F_i = S_i / I_t
\]  

(5)

where \( I_t \) is the monthly median income for lot \( i \), (as estimated by the census tract median income for families with children). Note that lots with no nearby ECE providers (i.e., with an \( A_i \) score of zero) will not have an affordability index score.

**QUALITY INDEX**

The quality access index was determined by the availability-weighted average quality score of nearby ECE services. Because Hawai‘i has no quality rating improvement system (QRIS) and does not collect systematic data on program quality, our options for quality data were limited. We defined quality as a dichotomous variable (scored 0 vs. 1). Providers with a national early childhood accreditation (NAEYC,
NECPA, NAFCC) and public preK classrooms were considered to be high quality. For each lot location \( i \), we calculated the quality index \( Q_i \) as the availability-adjusted average of the quality scores of all nearby providers:

\[
Q_i = \sum_{j \in U_i} (V_j \ast R_j / \sum_{j \in U_i} R_j)
\]

where \( V_j \) is the quality score of provider \( j \), \( R_j \) is the capacity-to-population ratio within the \( D \)-mile (or 30-minute) catchment area centered at provider \( j \), and \( U_i = \{ j : d(i,j) \leq D \} \) is a set of providers \( j \) that are within a \( D \)-mile (or 30-minute) commute. Lots with no nearby ECE providers (i.e., with an \( A_i \) score of zero) will not have a quality index score.

AGGREGATING INDEXES AT THE CENSUS TRACT LEVEL

For each census tract \( t \), aggregated index \( Index^T_t \) was summarized as a population-weighted average of lot level index scores, \( Index_i \):

\[
Index^T_t = \sum_{j \in t \& j \in U_j} Index_i * P_i / P_t
\]

where \( P_i \) is the number of children under age 6 at residential lot \( i \), \( P_t \) is the total number of children under age 6 at census tract \( t \), and \( U_j = \{ i : d(i,j) \leq D \} \) is a set of residential lots \( i \) that are in the \( D \)-mile (or 30-minute) catchment area centered at a given provider \( j \).

COMBINED INDEXES

Combined index scores represented the overall accessibility for each census tract. Unweighted and weighted versions of the combined index \( O_{iw} \) (\( O_{iw} \)) were calculated by taking the unweighted-average (or weighted average) of the Z-scores of spatial access \( A_i \), affordability \( F_i \) and quality \( Q_i \) as follows:

\[
O_{iw} = (Z_i^A + (-1) \times Z_i^F + Z_i^Q) / 3
\]

\[
O_{iw} = (2 \times Z_i^A + (-3) \times Z_i^F + Z_i^Q) / 6
\]

where \( Z_i^k = (k_i - K) / S_k \); \( k_i \) is spatial access \( A_i \), affordability \( F_i \), or quality access \( Q_i \) for each residential lot \( i \); \( K \) is the average of spatial access \( A_i \), affordability \( F_i \) or quality access \( Q_i \) for all lots the state; and \( S_k \) is the standard deviation of spatial access \( A_i \), affordability \( F_i \), or quality access \( Q_i \) for all lots. Higher values on the spatial access \( A_i \) and quality indexes \( Q_i \) indicate a more desirable level of access, while low values on the affordability index \( F_i \) are desirable; thus, we reversed the values of the affordability index in the equations (8a, b).

VISUALIZING AN OVERLAY OF THREE INDEXES

While the combined index has the advantage of showing spatial access, affordability, and quality as a single score, it has the drawback that tracts can have the same combined score based on different component scores. A visual overlay allows users to consider all three single indexes simultaneously, revealing an area's pattern of strengths or weaknesses across the three dimensions of ECE access.

---

5 National Association for the Education of Young Children, National Early Childhood Program Accreditation, and National Association for Family Child Care. Programs with any of these three national accreditations receive a higher subsidy reimbursement rate. Hawai‘i is among a small number of states to achieve the highest possible quality rating for its public preK program from the National Institute for Early Education Research.

6 The weights used in equation 8b were suggested by our stakeholder advisory committee.
To visualize disparities in accessing ECE services, we created a tract-level map of the overall accessibility index $O_i$ per residential lot $i$ by overlapping the spatial access $A_i$, affordability $F_i$, and quality $Q_i$ indexes as follows:

$$O_i = (A_i, F_i, Q_i)$$

We used two-color gradients and one pattern to distinguish the three dimensions.

**CREATING MAPS**

We used ArcGIS and ArcGIS online to calculate indexes and create a dashboard of interactive online maps. Each map contains a base map, a set of data layers with interactive pop-up windows providing detailed information about the data, and navigation tools to pan and zoom. We chose to use base maps that are hosted and shared through ArcGIS Online. Once the dashboards were assembled, we embedded them into a public website.

The website [home page](https://www.digmap.com/platform/smartparcels/) includes statewide summary statistics and a navigation video. A [provider map](https://www.digmap.com/platform/smartparcels/) shows the name, location, and licensed capacity of each ECE provider. To ensure privacy, individual information was not included for family child care homes. There are separate tabs for the nearby seats, affordability, quality, and combined indexes, where the user can then choose between tract- and lot-level views, and select the desired catchment area. Pop-up windows provide tract-level scores and demographic data, or lot-level scores, respectively.

**REFERENCES**


ESRI. (n.d.). ArcGIS Network Analyst extension. [Website].


SUGGESTED CITATION AND ACKNOWLEDGMENTS


This project was supported by a grant from the Robert Wood Johnson Foundation and by the National Institute of Food and Agriculture Hatch project 1017907. We are grateful to our sponsors for making this work possible.