Pilot Study:

Application of the Index of Child Care Access among five Early Learning Coalitions in Florida

PRESENTED TO:

The Florida Office of Early Learning

Greg Allerellie, Sunshine Donaldson

PREPARED BY:

The University of Florida Childhood Needs Assessment Partnership

Anita Zucker Center for Excellence in Early Childhood Studies

Herman T. Knopf, Phillip Sherlock, Sherry Zhou

UF Family Data Center

Mildred Maldonado-Molina, Roland Estrella, Lindsey Varner, Jiawei Li, Erik Schmidt, Idania Sanchez

Contents

Prepared for the Florida Office of Early Learning by The University of Florida Childhood Needs Assessment Partnership – OEL Contract # SR660 – July, 2018

Summary

Through this project the UF Childhood Needs Assessment Partnership endeavored to test the feasibility of applying the Index of Child Care Access in the state of Florida using only state administrative data. Due to the fact that the Florida SR program is locally administered, it was imperative to establish partnerships with Early Learning Coalitions to ensure that this research was informed by the wisdom and experience of local child care system administrators as well as responsive to the data and information needs of these skilled and savvy policy makers. Once these partnerships were formed, the expanded team embarked on this critically important journey to develop systems and build the infrastructure necessary to adequately measure the extent to which families have reasonable access to developmentally beneficial early care and education services that meet their needs. This collaborative effort resulted in the following notable outcomes:

1. Expanded the UF Childhood Needs Assessment Partnership to include 5 Early Learning Coalitions with fully executed data sharing agreements:
    a. Alachua County Early Learning Coalition
    b. Big Bend Region Early Learning Coalition
    c. Lake County Early Learning Coalition
    d. Pinellas County Early Learning Coalition
    e. Southwest Florida Region Early Learning Coalition
2. Developed a secured computing infrastructure with incorporated administrative and regulatory processes for data sharing of confidential data
3. Developed data intake, curation, and integration processes for OEL data
4. Developed a data codebook with updates specifications for data indicators
5. Developed a rationale and a proposed methodology for expansion of the Index of Child Care Assess to statewide implementation
6. Developed localized data reports (maps) illustrating application of the Index for Child Care Access combined with other pertinent administrative data for 5 Early Learning Coalitions.

The report that follows describes the processes, research methods, and products that have been developed through this new and innovative partnership among researchers and administrators from state and local levels. Our research and development efforts have been fruitful. We have indeed confirmed that the Index can be reasonably applied to inform our understanding of child care accessibility at the local level. This work represents a true breakthrough in our understanding of the importance of partnership between policy makers and researchers, that if maintained has the potential to enhance statewide capacity for data informed decision-making.

Introduction

The purpose of this pilot project is to use state child care administrative data to understand the patterns of subsidy use (i.e. location, provider type, and quality) among low-income families in select early learning coalitions in Florida. These patterns of use are identified through the refinement and application of the Index of Child Care Accessibility (hereafter referred to as 'the Index') which was initially developed by Knopf, Rao, Tester & Sherlock (2016) through the *Child Care Accessibility Index: Leveraging SC Child Care Administrative Data to Inform State CCDBF Subsidy Policies* with funding from the U.S. Administration for Children and Families, Office of Planning, Research and Evaluation (Grant Number: 90YE0176-01-00). The primary aim for this use of the Index is to support state (OEL) and local (ELC) child care administrators in making informed policy decisions to increase the use of high quality care among subsidy recipients. In service of this aim, the University of Florida (UF) Childhood Needs Assessment Partnership, a collaboration between the UF Family Data Center (FDC), the UF Anita Zucker Center for Excellence in Early Childhood Studies (AZC) developed a robust foundation for program quality and accessibility analysis in service to OEL. The new developed analytical capacity includes: 1) an expansion of the UF Childhood Needs Assessment Partnership to five ELCs (Alachua, Big Bend, Pinellas, Lake, SW FL) with fully executed data sharing agreements; 2) a secured computing infrastructure with incorporated administrative and regulatory processes for sharing of confidential data; 3) data intake, curation, and integration processes for OEL data; 4) a data codebook with updated specifications for data indicators; 5) pilot localized data reports (maps) illustrating application of the Index for Child Care Access combined with other pertinent administrative data for five ELCs; and 6) describe rationale and proposed methodology for expansion of the Index for Child Care Access for Statewide implementation.

Child Care Subsidy Helps Families and Children

Child care is an essential service that parents need to maintain the health and safety of their children while they are away from home earning a living or attending an education program to support career readiness. In Florida, the School Readiness Voucher Program helps more than 200,000 children from low-income families access needed child care services. Currently, throughout the U.S, and Florida in particular, there are several child care service provider options that parents can select from. These child care options include a provider type ranging from low/no cost options such as Family, Friend, and Neighbor care (FFN), to more moderately priced Family Child Care Home (FCCH), to the higher priced formal early care and education service providers such as child care centers; parents have many options to consider. When considering enrollment options, in addition to type of provider, parents consider characteristics such as location, price of care, quality of teacher child interaction, formal vs. informal setting, hours of operation, and type of curriculum (Blau & Robins, 1998; Blau & Tekin 2007; Blau & Tekin 2007; Chin & Phillips, 2001; Kimmel 1998; Tekin 2005, Tekin 2007; Weber & Grobe, 2011).

Influence of Child Care Subsidy on Family Employment

Since the federal subsidy programs have been consolidated within the CCDF as a result of the 1996 Personal Responsibility and Work Opportunity Reconciliation Act, researchers have

worked to determine the impact that child care subsidies have had on the child care purchasing power of families, subsequent employment, and self-sufficiency. Researchers investigating the importance of child care as a meaningful support for parental employment have documented that families who lack stable access to child care experience difficulties in securing and maintaining employment (Baum 2002). Furthermore, researchers examining the impact of subsidy receipt on parental employment have consistently found that low-income parents with subsidies are significantly more likely to work than similar low-income families without subsidies (Bambridge, Meyers, & Walfogel, 2003; Blau & Tekin, 2007; Brooks, Reisler, Hamilton, & Nackerud, 2002; Herbst, 2008; Tekin 2005). Research has also demonstrated that child care subsidies help families maintain employment as well as support increases in earnings (Bambridge, Meyers, & Walfogel, 2003; Brooks, Reisler, Hamilton, & Nackerud, 2002; Davis and Weber 2001; Forry, & Hofferth, 2011; Grobe, Weber, and Davis 2008; Ha, 2009; Meyers, Peck, et al. 2002; Tekin 2005) thereby increasing the likelihood that the cycle of poverty might be broken within these households. While subsidy receipt in and of itself is an important support for families, the type of provider that families select is just as important. Specifically related to supports for employment, research indicates that families that select more formal care arrangements, such as center-based providers, experience lower work disruptions and thereby experience increased job retention (Gordon, Kaestner, & Korenman, 2008; Knox, London, Scott, & Blank, 2003). While the selection of formal care arrangements is likely to support family employment, as described in the next section, evidence suggests that this type of care is also more supportive of children's growth and development.

Quality Early Care and Education Experiences are Critical

Within the field of early care and education, research evidence clearly points to the importance of high quality experiences for helping young children, especially those from low-income backgrounds, achieve positive social and academic outcomes (e.g., Burger, 2010; Crosby, Gennetian, Huston, 2005; Gormley & Gayer, 2005; NICHD & Duncan, 2003; Reynolds, A. J., Magnuson, K. & Ou, S. 2010). Investigations related to the associations of the type of childcare (i.e. home-based vs. center-based care) and the quality of care have found that children enrolled in high quality center-based care versus informal home-based care are more likely to demonstrate greater social-emotional and cognitive outcomes (NICHD & Duncan 2003; Forry, Davis, & Welti, 2013). The quality of child care necessary to support the above referenced positive child outcomes can be characterized by responsive teacher-child interactions, appropriately trained personnel, maintenance of basic provisions for health and safety, and adequate provision of equipment and materials to support learning. A number of early childhood investigators have determined that experimental (e.g., Lazar & Darlington, 1978; Schweinhart, 2005); state-funded pre-kindergartens (e.g., National Center for Early Development and Learning's (NCEDL) Multi-State Study of Pre-Kindergarten (2001-2002); the NCEDL-NIEER State-Wide Early Education Programs Study (SWEEP Study 2003-2004) see Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Downer et al., 2007; La Paro, Pianta, & Stuhlman, 2004; LoCasale-Crouch et al., 2007); and childcare programs (e.g., NICHD Early Child Care Research Network, 2000; NICHD Early Child Care Research Network, 2002) are associated with short-term developmental and learning improvements. Both narratives (cf. Farron, 2000;

Wat, 2010) and meta-analyses (cf. Camilli, Vargas, Ryan, & Barnett, 2010; Gilliam & Zigler, 2000) have confirmed these consistent positive findings with early childhood services improving child outcomes. Unfortunately, high quality care is not universally accessible for low income families (Coley, Li-Grining, & Chase-Lansdale, 2006; Dearing, McCartney, & Taylor, 2009; Torquati, Raikes, Huddleson-Casas, Bovaird, & Harris, 2011). Several studies in fact have demonstrated that while subsidy increased use of more formal child care arrangements, subsidy recipients tend to enroll in care characterized as poor to mediocre (Antle, Frey, Barbee, Frey, Grisham-Brown, & Cox, 2008; Jones-Branch, Torquati, Raikes, & Edwards, 2004). This level of quality is not likely supportive of positive child outcomes, as determined by recent research which indicates a threshold effect of quality necessary to achieve positive outcomes (Burchinal, Vandergrift, Pianta, & Mashburn (2010).

Using research innovations to strengthen data informed decisions: Moving from prediction to direct measurement

Limitations of sampling for nationally representative surveys

Most of the research on child care subsidies has used data from large national surveys such as Fragile Families and Child Wellbeing Study (FFCWS) Child Care Supplement (CSS), Early Childhood Longitudinal Study – Birth Cohort (ECLS-B), Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K), Survey of Income and Program Participation (SIPP), National Household Education Survey (NHES)'s Early Childhood Programs Participation (ECPP) and Before and After-School Program Activities (ASPA) of 2001 and 2005. While these national surveys provide sweeping insights, the CCDF child care subsidy program is a federal initiative with broad guidelines to be implemented at the state level that results in diverse program development. As a result, these national surveys often fail to capture the uniqueness of state child care subsidy programs which makes generalizing the results to inform policy a challenge.

Addressing the challenge of measuring child care accessibility

One of the primary goals of the CCDBG Act of 2014 is to help working families with low-incomes utilize child care subsidy to access high quality child care and therefore improve the development of young participating children (CCDBG Act Sec. 658A). A portion of CCDBG resources are, in fact, allocated to developing campaigns for parents to educate them on choosing high quality care (Sec. 658E(c)(3)(B), §9858c(c)(3)(B)). Parental choice in selecting a high quality child care provider is based on the assumption that there is availability of high quality child care and parents prioritize quality (Ryan et. al, 2011). Availability of child care is central to issues of choice and previous research on parent's child care choices (Henly, Ananat, & Danziger, 2006; Blau & Hagy, 1998). Due to the complexity of accessing and appropriately using administrative data for research purposes, virtually all of the research to date has studied parental choice of child care independent of availability within a proximity of the parents' residence. This pilot research project has allowed us to study child care use among child care subsidy recipients in the context of availability within zip codes among the 5 pilot Early Learning Coalitions. This work provides insights on 'revealed preferences' of child care by contextualizing

use of care by availability to parents and looking into patterns of child care choice, *when* they had a choice.

Through this pilot project, we analyze geographical regions with maps while taking into account the changing landscape of child care availability in Florida. This is necessary because child care utilization is a local phenomenon where local contexts, differ by demographics and geography. Administrative data are powerful because they reflect local policies, procedures and practices, the relevance of which cannot be overstated as a tool for guiding the decisions of state child care administrators. This project fills gaps in local and state knowledge while basing the theoretical framework of the study on previous research done in the national context using secondary data and national surveys.

Addressing existing limitations through intentional partnership

Guided by the mission of applying what we know about measuring family access to develop a more refined and geographically relevant tool to inform critical policy making decisions, the UF Childhood Needs Assessment Partnership recruited 5 Early Learning Coalitions (ELC of Alachua County, ELC of Big Bend Region, ELC of Lake County, ELC of Pinellas, and ELC of Southwest Florida) to participate in the pilot project. Providers were selected to participate based on recommendations provided by the FL Office of Early Learning, confirmation of interest, and variation in demographic characteristics. Voluntary participation in the pilot project expected a commitment from the Early Learning Coalition representatives to comprehensive data sharing and use agreements and engage with the research team through monthly advisory committee meetings, wherein the advisory committee brainstormed policy questions, helped to define "developmentally beneficial early care and education services" and help provide social validity to the process and products that have been developed through this project. In addition to the Advisory Committee meetings, ELC partnering agencies were consulted individually regarding more specific regional data accuracy checks, social validity checks of draft data visualizations, and provided input regarding the interpretability and utility of our final mapping reports.

Index of Child Care Access

Child care is a local phenomenon with community markets differing based on demographics and geography. Unfortunately, children in many communities do not have reasonable access to the type and quality of services necessary to develop the knowledge, skills, and dispositions necessary for school and social success. Through the reauthorization of the Child Care Development Fund in 2014, the United States federal government instituted new regulations that specifically direct states to make strategic investments to increase child and family access to high quality early care and education services (CCDBG Act Sec. 658A). The Index of Child Care Accessibility was formulated in response to that regulation. Requiring only administrative data, the Index is a systematic tool for detecting differential access to child care for subsidy recipients. It can be used by any local or state government to guide policy development. The Index describes the state of child care markets in different geographic areas using two sub-indices (selection and infrastructure) for a specific point in time. The sub-indices serve as proxy measures for understanding child care utilization. Each of the two sub-indices can be either

negative or positive, resulting in four possible combinations of accessibility for a given geographic area. After calculating each of the index scores for a given geographic area, administrators and researchers can use the Index Policy Matrix to identify areas that warrant deeper, more localized analysis to develop logical, locally-informed interventions.

The Index Sub-Indices

In Florida, as is the case for many states, there is no uniform practice for obtaining child care vacancy information. Considering this, our research team constructed two proxy measures (*selection* and *infrastructure*), representing two sub-indices. Both the selection and infrastructure sub-indices incorporate observed measures at the zip code level: (1) the type of care facility; (2) the number of children using vouchers in those types of facilities; (3) the total capacity of the settings; (4) the quality levels of the facilities. The first sub-index— selection— captures the selections made by parents within the context of other possible selections. The second sub-index—infrastructure— indicates the degree to which all of the children receiving subsidies in a given zip code *could* attend high quality center-based care.

Selection

The selection sub-index shows how well parents are making decisions to enroll their children in the highest quality care given the context of availability in their respective zip codes. More specifically, selection measures the difference in the proportions of subsidy recipients' enrolling in gold seal providers (high quality) versus non-gold seal providers given zip code capacity at the respective types of providers. Selection can be calculated for zip codes that have provider capacity greater than zero and at least 1 child served in a contracted SR provider. Weights for the different quality levels have been incorporated in the formula to place importance on the parental enrollment decision. Gold Seal providers are given a group weight equal to positive one. All other centers (i.e., non-Gold Seal providers) are given a group weight equal to negative one. Table 1 shows the weighting scheme using Florida's Gold Seal status as a proxy for high quality.

Table 1. Example Quality Groups and Weights

Quality Group	Gold Seal	Non-Gold Seal
Weight	+1	-1

A positive selection value, X, indicates the proportion of children enrolled in Gold Seal care divided by the Gold Seal care capacity was greater than the proportion of children enrolled in Non-gold Seal care divided by the Non-gold Seal capacity. A negative selection value, -X, indicates that the proportion of children enrolled in Non-gold Seal divided by the Non-gold seal capacity was greater than the proportion of children enrolled in Gold Seal divided by the Gold Seal capacity.

Infrastructure

Infrastructure is used to identify whether a zip code has an abundance or shortage of high quality care relative to the number of children using subsidized care within the given zip code.

Infrastructure was calculated for zip codes with one or both of the following: (1) at least 1 child served in subsidized care; (2) high quality capacity. Infrastructure measures high quality capacity minus the number of children receiving subsidies in a zip code.

A positive infrastructure value, Y, indicates how many more Gold Seal slots than children receiving subsidies. A negative infrastructure value, -Y, indicates how many more children are receiving subsidies than there are Gold Seal slots.

Measuring Access in the spatiotemporal context

The Index of Child Care Access is intended to be a direct measure of how well families access available child care services that meet their needs and support children's development and learning. As a direct measure of the child care selection and use behavior of families using SR vouchers, the Index intentionally calculates selection and infrastructure for a particular point in time. By including the time component to the application of the Index, our team is able to measure change in both parent selection and infrastructure. This spatiotemporal measurement increases the utility of the Index to reflect the impact of system and policy changes through time, helping policy makers evaluate the effectiveness of their system in achieving the goals if increasing access to high quality services for children.

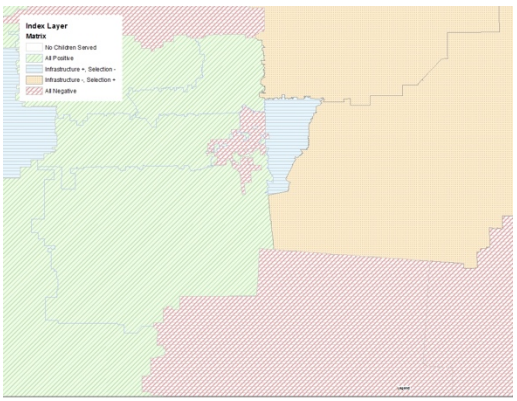
The Index of Child Care Access Policy Matrix

The Index of Child Care Accessibility Policy Matrix is a tool for interpreting the interplay of the two Index sub-indices. The Policy Matrix has four cells that refer to four different possible scenarios for a geographic area: (1) negative selection, negative infrastructure; (2) positive selection; negative infrastructure; (3) negative selection, positive infrastructure; (4) positive selection, positive infrastructure. The Index Policy matrix is used by policymakers to identify the types of interventions necessary for increasing access to high quality care in a zip code for families receiving subsidies. Each cell contains policy recommendations based on selection and infrastructure in that zip code. Figure 1 displays The Index Policy Matrix with the negative/positive selection sub-index on the top row and the negative/positive infrastructure sub-index on the left-hand column. Figure 2 displays the zip code level map where the color of each zip code refers to a block in The Index Policy Matrix based on the particular combination of the selection and infrastructure sub-indices. White areas on the map correspond to areas in which no subsidy recipients were served.

Figure 1: Policy Matrix

	Selection	
Infrastructure	Positive (+)	Negative(-)
Positive (+)		
Negative (-)		

Figure 2 Example of Index visualization



Scenario 1—Positive Selection and Positive Infrastructure

In the first scenario, the upper-left corner of The Index Policy Matrix, both the selection and infrastructure sub-indices are positive. This means that in the context of the SR contracted care available to parents, a greater proportion of parents are choosing high quality. Specifically, there are a greater number of high quality child care slots than there are children receiving subsidies in these areas. Based on the fact that both sub-indices are positive, administrators and researchers should investigate these areas to better understand how to implement interventions in the other three types of areas.

Scenario 2—Negative Selection and Positive Infrastructure

In the second scenario, the upper-right corner of the Index Policy Matrix, the selection sub-index is negative and the infrastructure sub-index is positive. This means that in the context of the SR care available to parents, a greater proportion of parents are choosing Non-gold seal care. However, there are a greater number of Gold Seal child care slots than there are children receiving subsidies. Given the poor selection by parents despite there being more Gold Seal slots than children receiving subsidies, The Index Policy Matrix indicates a need for administrators and researchers to investigate why parents are not choosing Gold Seal providers and develop interventions to improve parent selection (e.g., a parental awareness campaign).

Scenario 3—Positive Selection and Negative Infrastructure

In the third scenario, the bottom-left corner of the Index Policy Matrix, the selection sub-index is positive and the infrastructure sub-index is negative. This means that given the SR care available to parents, a greater proportion of parents are choosing to use it. It also means that there are a greater number of children receiving subsidies than there are Gold Seal slots. Given that parents are more often enrolling their children in Gold Seal care despite the shortage of Gold Seal infrastructure, The Index Policy Matrix indicates a need for an increase in Gold Seal slots to make sure that all children receiving subsidies are able to enroll in high quality care.

Scenario 4—Negative Selection and Negative Infrastructure

In the fourth scenario, the bottom-right corner of the Index Policy Matrix, both the selection and infrastructure sub-indices are negative. This means that given the SR care available to parents, a greater proportion of parents are choosing low quality care. It also means that there are a greater number of children receiving subsidies than there Gold Seal child care slots. Based on the negative selection sub-index, the Index Policy Matrix indicates a need for administrators and researchers to investigate why parents are not choosing the Gold Seal providers and develop intervention to improve parent selection (e.g., a parental awareness campaign). Additionally, based on the negative infrastructure sub-index, there is a need for an increase in Gold Seal slots to make sure that all children receiving subsidies are able to enroll in Gold Seal care.

Data Collection and Analysis

To accomplish the development of The Index, it was crucial to create a custom-built process of data security, intake, storage, curation, and output to ensure compliance with national

standards for human data privacy and security. The following sections outline a total of eight processes developed and implemented by FDC to accomplish the following two goals:

7. A data security network that simultaneously prioritizes the security of sensitive data and efficiency of data processing;
8. A data curation process that is precise, thorough, and efficient.

FDC Data Security Network

The OEL-Index of Child Care Access project utilizes a highly-secured server infrastructure that was custom-built by the UF's Research Computing. The system was developed based on systematic testing and feedback from the UF Childhood Needs Assessment partnership faculty and staff who have expertise in technical architecture as well as data domain knowledge. Based on testing during the pilot phase of the project, the developed environment shows efficient performance and guaranteed uptime, which is a core requirement for the Index workflow. The following paragraphs describe the ResShield infrastructure and outline the information flow through data intake, access and authentication, data curation and aliases, and limited data sharing as shown in the system architecture diagram (see Figure 3).

Data Intake

All data received from OEL are moved via secure file transfer (SFTP) directly into the UF ResShield environment. From there, the partnership data custodians are able to verify receipt and completeness. After verification, data is then transferred by data custodians into a virtual DataBank server on which the necessary software tools have been installed for data curation, including Microsoft SQL, SAS, ArcGIS, and R. The data transfer is performed using a secure and encrypted tool, which is configured specifically for OEL data transfers.

Authentication

The ResShield infrastructure meets the highest security standards because it employs network firewalls and state-of-the-art protections that are designed to limit access to only one single connection point. This connection is through a Virtual Desktop Interface (VDI), which requires users to be on a restricted network with multiple (interconnected) levels of security including Virtual Private Network connections (VPN). Once access to the VDI is gained through VPN, the authentication process is multi-factor requiring users to confirm their intent to login through an encrypted application on their mobile phone. These multiple layers of protection help ensure that project data, and all access to the project data, are isolated from any other project or network traffic internally, even internally at UF. The environment has been FISMA certified in accordance with NIST 800-53, a federal data security standard.

Data Curation and Analyses

Once data is secured inside the ResShield virtual DataBank environment, authorized project analysts conduct a four-stage data curation process, which prepares data for analysis (e.g., the extract, transform, load (ETL) process described in Section 2). The data curation workflow involves stripping of identifiers that may not be transmitted to downstream analysts. Once data is curated and transformed into output datasets, the partnership data custodians move this transformed, limited data to a separate virtual server within ResShield developed for the sole purpose of the Index analyses with limited data. In this further isolated ResShield environment,

the the Index faculty and staff can conduct data analysis and develop reports, but strict data protections still limit traffic in-and-out of the the Index work environment. Once aggregated data and reports are ready to be shared with authorized partners and OEL, a request must be made to the partnership's data custodians so that files are copied out of the the Index environment and into the UF campus network. The UF campus network is still encrypted and secured, but has external connectivity including access to email and the ability to do videoconferencing with external entities.

Sharing of Limited Data

In sharing aggregated data with external entities, the UF partnership also took effort in deploying tools that ensure data is protected at all times. To guarantee that even the aggregated and/or limited data remains secured, a special edition of Zoom videoconferencing software configured for restricted information was made available to OEL and the Index project faculty and staff.

Barriers, Lessons Learned and Reflections

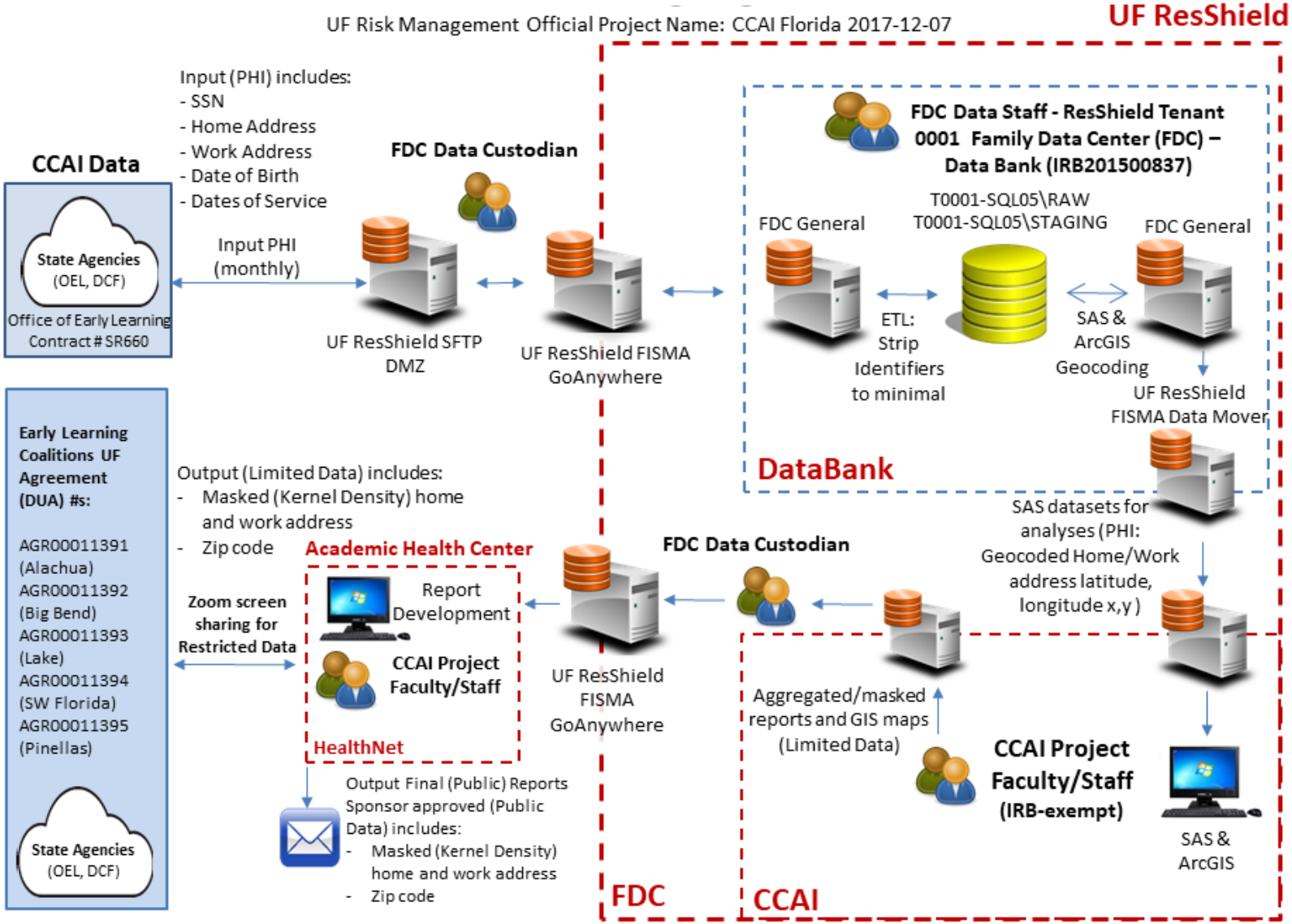
The challenges encountered throughout the development of a custom process of data security and data curation provided insight to goals for the statewide expansion of The Index application. On reflection, the two key processes in data processing (data intake and transfer security and data curation) are naturally at odds. Our completion of this pilot phase, which was able to strike a balance between security and speed of processing, is encouraging for future expansion. Listed below are barriers and associated recommendations for statewide scaling.

While the security infrastructure developed for the pilot project meets strict federal government security standards, naturally, it decreases the efficiency of the data curation processes due to manual vs. automated operational work-flow elements. For example, if at any point during analyses a data transfer was needed (in or out of the secured environment), an authorized data custodian had to manually inspect the granularity of the data and approve and conduct the transfer. This approval process relied on unanticipated scheduling constraints, which impacted moderately the analytical workflow. With these semi-automated data security and work-flow quality control processes in place, the data curation component of the project was satisfactory for processing data from the five pilot ELCs, however, for expansion, further automation is recommended with the goal of streamlining the secured data flow security processes, so that analytical elements of the work cycle remain efficient and transition (quickly) from input to output work-flow stages. The following bullets outline lessons learned from our tests of multiple distinct security environments and the corresponding reflections on expansion.

1. The infrastructure uptime is essential to project success. The Research team had multiple critical periods of work, which were impacted by downtime in one of the test environments.
    9. Reflection: Moving forward, we do not anticipate uptime to act as a barriers to project success. The ResShield system is a mature environment for data security and has record uptime and stability as compared other comparable, secure environments.

2. It is important that the Research Computing team in support of the network infrastructure respond to work stoppages with fast and accurate support.
    10. Reflection: The UF ResShield IT team is a powerful ally in both securing data and guaranteeing efficient work-flow in a production environment for statewide data feeds. During critical periods of analyses work, the UF ResShield IT team provided round-the-clock support and innovative troubleshooting. We have developed a strong partnership with the UF ResShield IT team, which we know will be instrumental to the project's statewide expansion success. Because of their capacity to provision secured, high-performance technology, operationalize strict regulatory compliance controls and processes, and provide outstanding customer service, we recommend that continue to operate and grow the CCAI UF ResShield environment with support from OEL.
3. The data intake aspect of the curation process must be conducted on a firm schedule for point-in-time analysis. Data collection depended heavily on contract amendments and data use agreements both needing administrative time.
    11. Reflection: Expanding to a statewide data collection and processing will benefit from a streamlined regulatory process that aligns Data Sharing Agreements (DSA) and contracting schedules with those of data intake and data transfers. The FDC maintains a day-to-day strong working relationship with the UF Division of Sponsored Research (DSP). This division is responsible at UF of processing legal documents including contract amendments and DSAs. Our recommendation is to engage with our partners in that office as well as with their OEL and DCF counterparts to work on operationalizing a work system to streamline creation, routing, and approval for regulatory documents in ways that those processes become aligned with the data intake and output work-flow of the UF Childhood Needs Assessment partnership to effectively integrate the Index within the ECENA data portal.

Figure 3. FDC Data Processing Diagram



FDC Data Curation Process

The University of Florida (UF) Childhood Needs Assessment Partnership designed and conducted a data curation process with the Office of Early Learning (OEL) to support the creation of the Index of Child Care Access Pilot. From within the secure FISMA environment, partnership researchers developed a data curation plan based on feedback from both technical and early education experts. The data curation process includes the following workflow components:

- Data warehousing
- Address geocoding
- Masking of restricted data
- Quality control

Data Warehousing

Database warehousing for this project focuses on data normalization, which is the process of restructuring data to reduce redundancy and improve integrity. A star schema is the simplest style for building dimensionality in normalized data and is the approach most widely used to develop data warehouses. The star schema separates business processes data into facts, which hold quantitative data about the business processes, and dimensions, which are descriptive attributes related to fact data. To normalize the data received from OEL, the partnership designed four individual dimensions at the top-most level of the data flow: Providers, Parents, Children, and Services. By separating these data elements into dimensions, the partnership was able to provide analysts with data structures that are able to uncover (by means of simple queries) how these four components are interconnected. The relationships among the four dimensions might seem simple for the purposes of descriptive statistics, but for the purposes of advanced analyses (e.g. continuous data exchange and geographical clustering techniques), having the ability to access the underlying complexities of why, when, and how often Providers are linked to Parents, Parents are linked to Children, and Services are provided to Children over time, offers the most benefits for extracting operational and strategic insight from the data.

By using SQL server integration services and SAS, the partnership developed a variation of the star schema approach (see Figure 4) to parse the data into one fact table and multiple dimensions, to develop the Index data warehouse that is simple, effective, and flexible. This approach, while initially intensive, normalizes the data (i.e., reduces redundancy and improves integrity) in a way that allows the UF partnership to streamline future recurring data curation processes.

Masking of restricted data

Once data has successfully been transformed into a star-schema structure, the next step on the data curation workflow is data masking. The data received from OEL includes Provider, Parent, and Children personal identifiers, which are confidential. Even though these identifiers are initially required to be able to create the star schema data structure (ie. connect the dimensions), moving forward in the data curation workflow, they are no longer meaningful to the analyses, so they can be masked. To mask these identifiers, the partnership uses a (key, value) pair system. Each value for each of the data elements to be masked is assigned a surrogate numerical ID that replaces the actual raw data on the star-schema with a numerical

key before the data moves through the curation process workflow. In other words, confidential data gets detached from the data that is passed-on to the project analyses and a random numerical key is put in its pace instead. This process effectively strips the data to be passed-on to the analytical workflow to the minimum necessary elements used for the analyses. Masking data tables containing key, value pairs are identified in blue in Figure 4. These tables are not used on to subsequent data curation steps.

Address geocoding

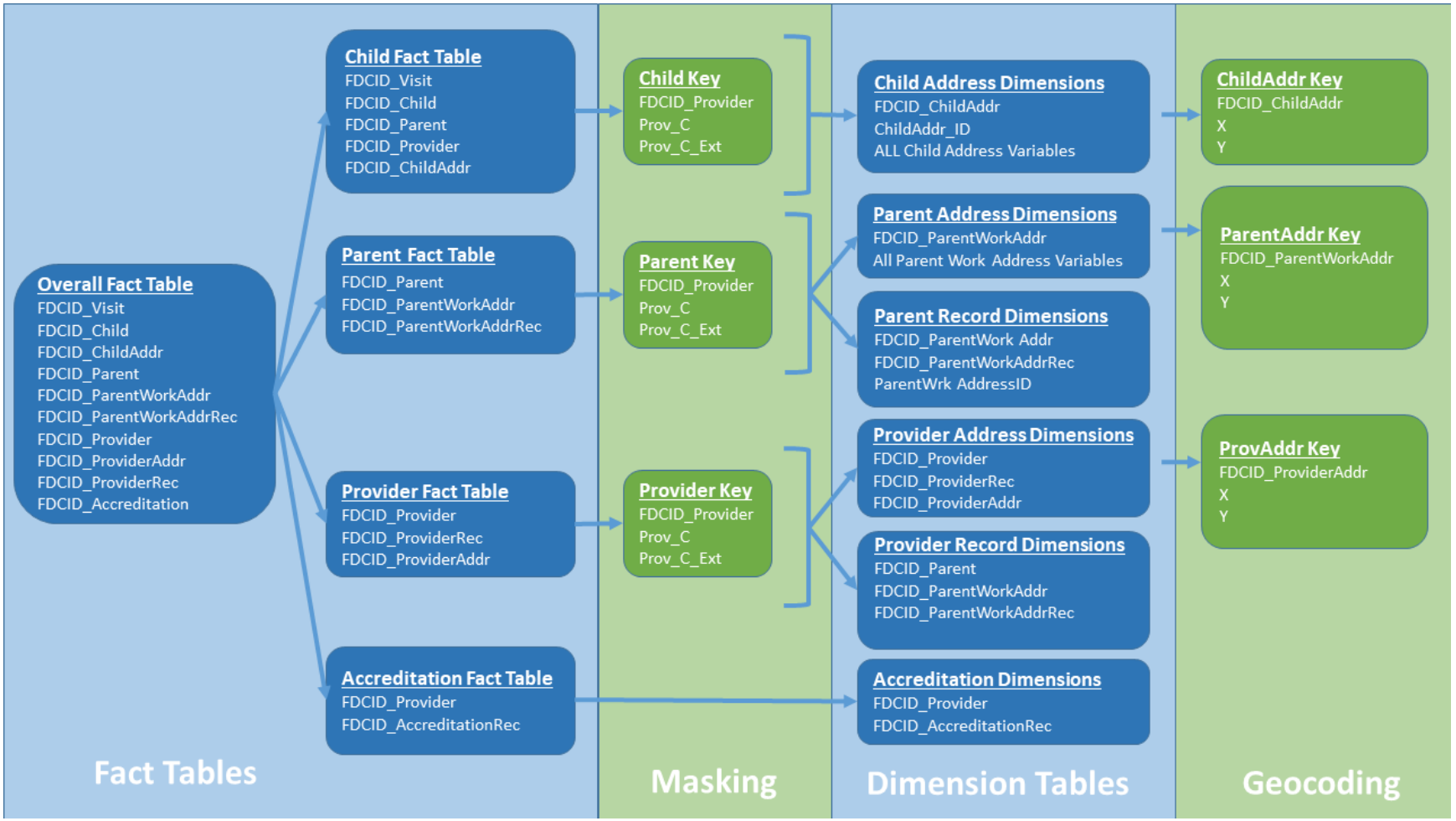
Once the restricted data has successfully been masked, the data curation process advances to the next stage, which is Geocoding. Geocoding is the process of transforming an address to a location on the earth's surface (latitude, longitude), which can be used for mapping or spatial analysis. Geocoding is an important part of the data curation process developed by the partnership. OEL transmits to UF address information for each Provider, Parent's work location, and Child's residence. By geocoding these addresses, the partnership is able to transform tabular data into spatial representation via maps. Maps can help inform decision makers at the State and local levels towards developing interventions more efficiently and strategically investing to increase accessibility of high quality services to vulnerable populations. Maps that use Geocoded data are particularly helpful in helping with identification of vulnerable populations as opposed to maps that publically available which are essentially pre-aggregated based on administrative boundaries such as counties, zip codes, or census tracts.

In order to Geocode the address data received from OEL within the off-grid computing environment, the partnership deployed the StreetMap Premium software to make the geocoded information available to project analysts. StreetMap Premium provides high-quality data, optimized for cartographic map display, geocoding, and routing, that works offline. The geographic locators installed with Street Premium are updated quarterly with the most recent and most accurate street data available worldwide, which is consolidated in the StreetMap Premium product, but it is collected by digital map data vendors including HERE. A technical overview and demonstration of the ArcGIS StreetMap premium product can be found in the link provided below: https://www.youtube.com/watch?v=dg3z8p1xWdA

Quality Control

The process of quality control consists of reconstructing the original OEL data file from the normalized, geocoded, and masked star-schema data structure. This reverse development process, once the partnership confirms that the output data is an exact copy of the initial (original) input data, serves to validate the consistency of the process, and ensures that all the relationships uncovered in the data structure, as well as all of the masking layer, are aligned with the original data. Lastly, deeper-level manual data checking is conducted to verify that the automated data validation was successful. The final product of the partnership data curation efforts are fully validated output files containing all child, parent, provider, and accreditation data provided to the Index pilot analytical team in their format of choice.

Figure 4. Family Data Center Star Schema for Child Care Access Data



Getting Acquainted with Florida's Administrative Data

Embarking on research that uses administrative data for a secondary purpose requires that the research team spend adequate time inspecting, validating and exploring the data to gain an intimate understanding of the information contained within the administrative data sets so that appropriate analyses are conducted and results are accurate and meaningful. Upon receiving the curated data from the FDC the AZC research team engaged in the following activities to acquaint themselves with the Florida OEL SR program data.

Develop a Data Codebook

The first task undertaken by the AZC research team upon receiving the curated data filed from the FDC, was to develop a data codebook (Appendix A). This document served as the initial step in support of the team getting acquainted with the data. Through this process the team communicated with the OEL data quality officer, FDC team members, and carefully inspected the data to ensure that accurate definitions were established and subsequently applied when engaging in further review and analysis. The result of this initial process was the development of a data codebook, sometimes referred to as a data dictionary that was consulted throughout the project.

Data validation and accuracy

Once the codebook was developed the AZC research team completed a thorough data review and quality check process. Through the processes of this stage, the team completed analyses of data completion, conducted data validation and accuracy checks by duplicating publicly available state School Readiness voucher disbursement reports, and shared initial findings with relevant state and Early Coalition representatives for feedback and to verify our understanding.

Findings from data inspection

The data inspection and accuracy checks revealed that by and large, the data obtained from the FL Office of Early Learning was indeed valid and accurate and complete, particularly as related to the child file, which reported SR voucher use. Inspection of the provider file from OEL revealed that the file only contained providers who had received a payment that coincided with the time period for which we had obtained child SR voucher use. Functionally, this means that the data did not include all possible providers from which families selected child care, it only included providers that were selected by families. This limitation is problematic when endeavoring to account for the entire choice set that parents had at the point in time that they were selecting a child care provider. In addition, this limitation suppresses the number of child care slots that would be used to calculate both the selection and infrastructure sub-indices. A second limitation of the provider file was that the provider capacity variable column was incomplete, wherein many providers did not have a capacity listed. This limitation also negatively impacts the accuracy of Index calculations as the total capacity within a geographic region is suppressed.

Engaging the Index Advisory Committee: Defining Quality, Collaborative Troubleshooting, and selecting a snapshot date

Subsequent to the initial data inspection and validation the UF Childhood Needs Assessment Partnership convened meetings with the Index Advisory Committee to obtain feedback on initial findings, discuss options to address limitations in the data obtained to that point, and

engage in discussions regarding a unified definition of high quality early care and education providers.

Defining Quality

A critical decision point that needed to be addressed by the Advisory Committee was establishing a minimum quality threshold represented by data indicators available at the local Early Learning Coalition. This quality designation is necessary to appropriately weight the capacity when calculating the Index. To establish this threshold the UF Childhood Needs Assessment Partnership shared a list of data that were available and discussed the strengths and limitations of each (i.e., Gold Seal Status, specific accreditation). Through these discussions the advisory committee agreed that it would be useful to designate Florida Gold Seal providers as meeting the threshold of high quality early care and education providers. While the Advisory Committee acknowledged that this is not a perfect proxy for high quality due to the plethora of recognized accreditation granting organizations and the variation in criteria and validation practices associated with same, this was *an* indicator that the provider is voluntarily providing services above and beyond merely meeting state child care regulations. Furthermore, this is a designation that is universally accepted throughout the state and therefore would apply to all pilot Early Learning Coalitions regardless of different local policies and practices related to locally administered quality measurement systems. As the legislative session progressed and HB 1091 was passed, initializing the state-wide implementation of the Classroom Assessment Scoring System (CLASS) among all SR contracted providers, the committee agreed to include CLASS as a measure of quality in future implementations of The Index. Once the CLASS is fully implemented throughout the state, its inclusion in The Index will allow for a more sophisticated model for quality designation.

Troubleshooting data limitations

In addition to providing excellent council in identifying a defensible definition of high quality early care and education providers, the Index Advisory Council was instrumental in helping to trouble shoot the provider data limitations. Through our deliberations the Advisory Committee was able to identify the publicly available DCF reporting of child care providers as a viable source of date reflecting the universe of legally operating child care providers throughout the state of Florida, but also likely complete regarding provider capacity, SR status, and Gold Seal Status.

Selecting a snapshot day

As described earlier, the Index of Child Care Access measures the geospatial characteristics of family selection within the context of available infrastructure. This is accomplished through the identification or a point in time relevant to policy decision makers. While the ultimate goal of this project is to implement a system to consistently measure child care access, the pilot project was designed to test the ability to apply the Index in a more locally administered child care system. Therefore, the AZC research team consulted the Index Advisory Committee to select an initial date for analysis. Based on their interest in reviewing complete and relatively recent data, the group selected June 2017 as the initial data for analysis.

Incorporating additional data

Following the Index Advisory Committee decision to pursue the acquisition of Provider level data from the Florida Department of Children and Families, the FDC team obtained an extract from the ECENA portal, appended geographical location coordinates, and transferred that data set to the AZC research team to include in our analysis.

Import provider level information from FL Department of Children and Families (Public Data)

After obtaining the DCF provider data files, the AZC team joined the DCF file with the existing OEL provider file to be used to calculate the infrastructure and selection indices. Once the data join was complete the AZC research team inspected the unified output file to verify successful and appropriate joins, and reviewed all cases where a provider from either file didn't match. In most instances, the non-matching providers could be explained through either program closure, program opening, or program termination from the SR program.

Caution: Misalignment of point in time limits the specificity of the pilot project outputs

This join/match analysis did reveal another complexity present in the data, the date that services were provided need to be taken into consideration when comparing family selection with available infrastructure. Our analysis of the match between the two data sets revealed that the child care industry in Florida is highly volatile with many providers initializing operations and ceasing operations each month. The merged data set, OEL child and provider files for services provided June 2017, was meaningfully different from the list of providers in the DCF data file obtained March 2018. This is a significant limitation in the data used in this pilot project, outputs from the analyses should not be used to inform policy decisions until this limitation is addressed. Fortunately, this is a limitation that is indeed addressed through a revised data collection process that establishes parallel acquisition of data from OEL and DCF.

Visualizing Child Care Accessibility: Using maps to inform policy making

The maps that our team has created are intended to demonstrate the capability of the Index of Child Care coupled with GIS technology to inform our understanding of the extent to which existing policies and practices work to facilitate increased access among families using child care subsidy to attend enriching early care and education programs. The sections that follow describe the different layers of the maps that we have developed highlighting the important features and how they contribute to our understanding of the early care and education system in the 5 pilot Early Learning Coalitions.

Base Layer: County/ELC Designations

The base layer of the map merely shows the county designations throughout the state of Florida. This layer of the map highlights the ELCs that were included in this pilot (white) and has green the other areas of the state. Zip designations of the map provide additional definition to the geography while also providing information that is useful for identifying the market within which families are choosing providers to care for their children. The grey lines in this map layer indicate zip code boundaries.

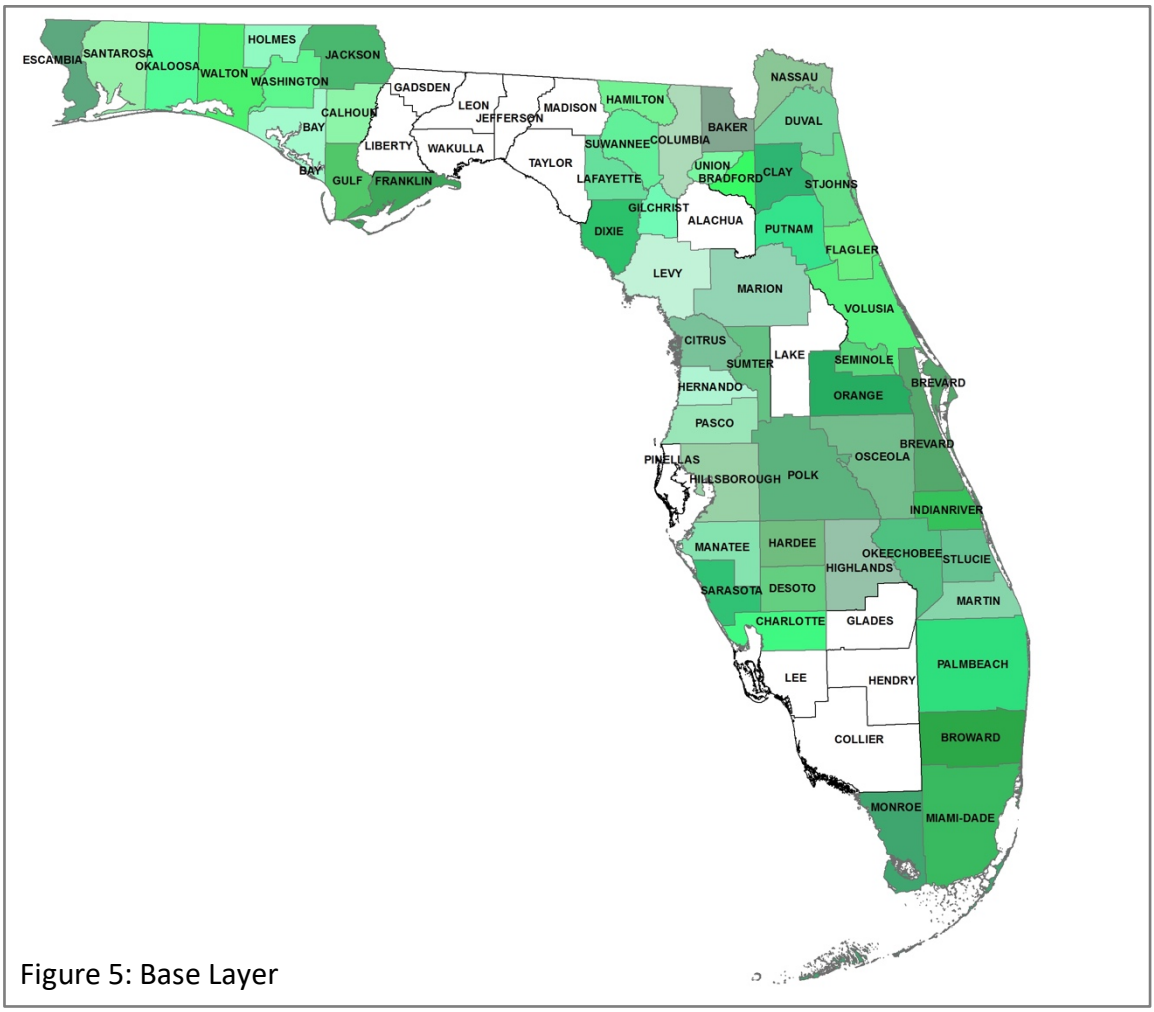


Figure 5: Base Layer

Index of Child Care Access: Policy Matrix

The Index layer of the map shows the output from the coalescence of the selection and infrastructure calculations for each of the zip codes included within the Early Learning Coalitions selected to participate in this pilot. The zip codes are color coded to correspond to the policy matrix to increase the interpretability and ease with which the local policy maker can identify the aspect of child care access that might warrant policy intervention.

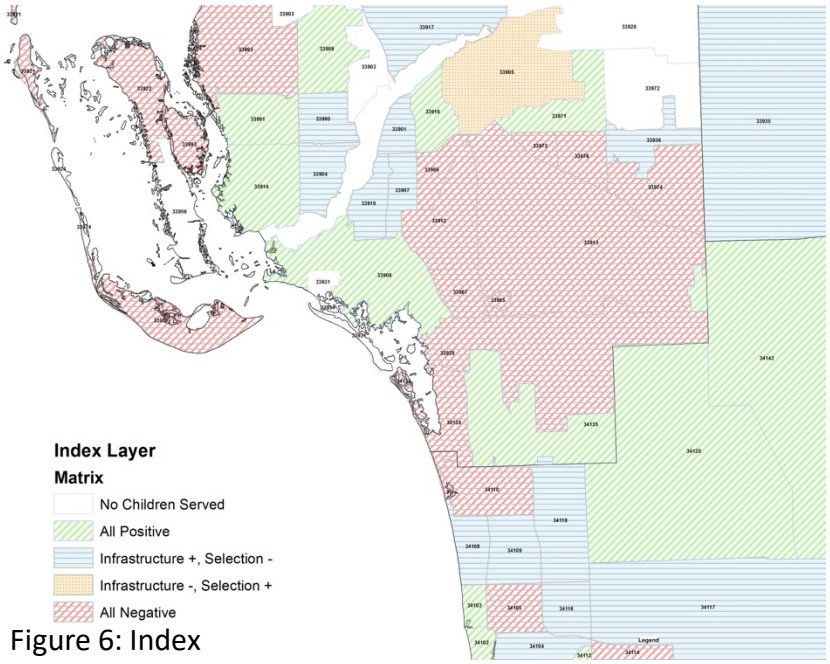


Figure 6: Index

In areas that are shaded in **green**, both family selection and existing infrastructure appear to be functioning well. The **blue** area of the map show areas where infrastructure is positive, but the selection, or families effective use of available Gold Seal child care slots, is less than optimal. These blue areas of the map are likely to be improved through a better understanding of the parent selection practices, or constraints that might be constraining their choice set. Some possible culprits include, high priced child care providers, provider-imposed limits of SR vouchers they are willing to enroll, provider hours that do not need the work and/or school demands of the family. The **orange** areas of the map show that while the family use of available Gold Seal slots is good, the area does not have adequate infrastructure to serve all families using SR vouchers within the given zip code. This situation is improved through an increase in available infrastructure, which in most cases is easier said than done. To increase infrastructure within a given zip code, the Early Learning Coalition can either identify Gold Seal providers not currently contracted to provide SR services and convince them to start, or work with existing programs to provide the professional development supports necessary to increase operational quality so that Gold Seal status is achieved. On the opposite end of the continuum, areas that are shaded **pink** indicate that both infrastructure and parent selection behaviors are not functioning well. This is a complicated situation to address, which would require additional information and analyses drawn from existing administrative as well as the collection of additional information from families using care in this area as well as providers serving families.

Family Residence Kernel Density Clustering Layer

The family residence kernel density clustering layer provides additional information that can be used to inform local administrator of any potential relationship between the home residence of children receiving SR vouchers to enroll in child care and the selection and infrastructure indices. This is particularly helpful when endeavoring to identify which areas, among several that might benefit from policy interventions, to focus initial investment and intervention. For example, when deciding between two pink zip codes (areas with negative infrastructure and selection) one might direct focus to areas with a relatively high SR population density, since interventions in those areas will benefit a greater proportion of families served by the program. Conversely, a local SR program administrator might also be interested in exploring a green area (positive selection, positive infrastructure) that also has a high SR population density, to better understand the characteristics of that community that are supporting a positive state of affairs.

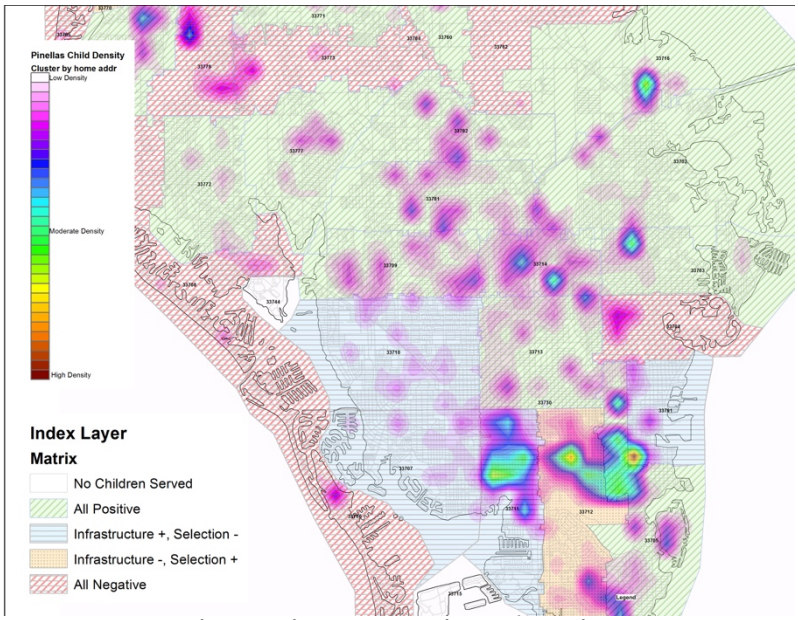


Figure 7: Family Residence Kernel Density Cluster

Provider Layer

The provider layer of the map pinpoints the location of child care providers. In addition to showing the location of the providers we have coded the providers by type, quality, status as a School Readiness Provider, and capacity. This information helps provide additional context with which to interpret the extent to which children and families have reasonable access to Gold Seal quality child care within each zip code of a community. This additional layer, when combined with the Index, helps to determine whether or not a particular course of action might be reasonable given the specific zip code. For example, the image in Figure 9 shows zip code 32726 which is blue, indicating good infrastructure, and negative selection. By focusing on this area, one can see that while there is adequate selection, there are far more non-gold seal providers in the area that could be influencing the choices that families make, impacting the price/affordability of the service when comparing gold seal to non-gold seal, or the gold seal providers in the area could be restricting the number of families using SR they will enroll in their center. While the policy implications for each of these scenarios are different, it is clear that the map helps the local policy maker refine her/his understanding of access in the area, and helps to identify additional questions that must be investigated to result in a truly data informed decision-making process.

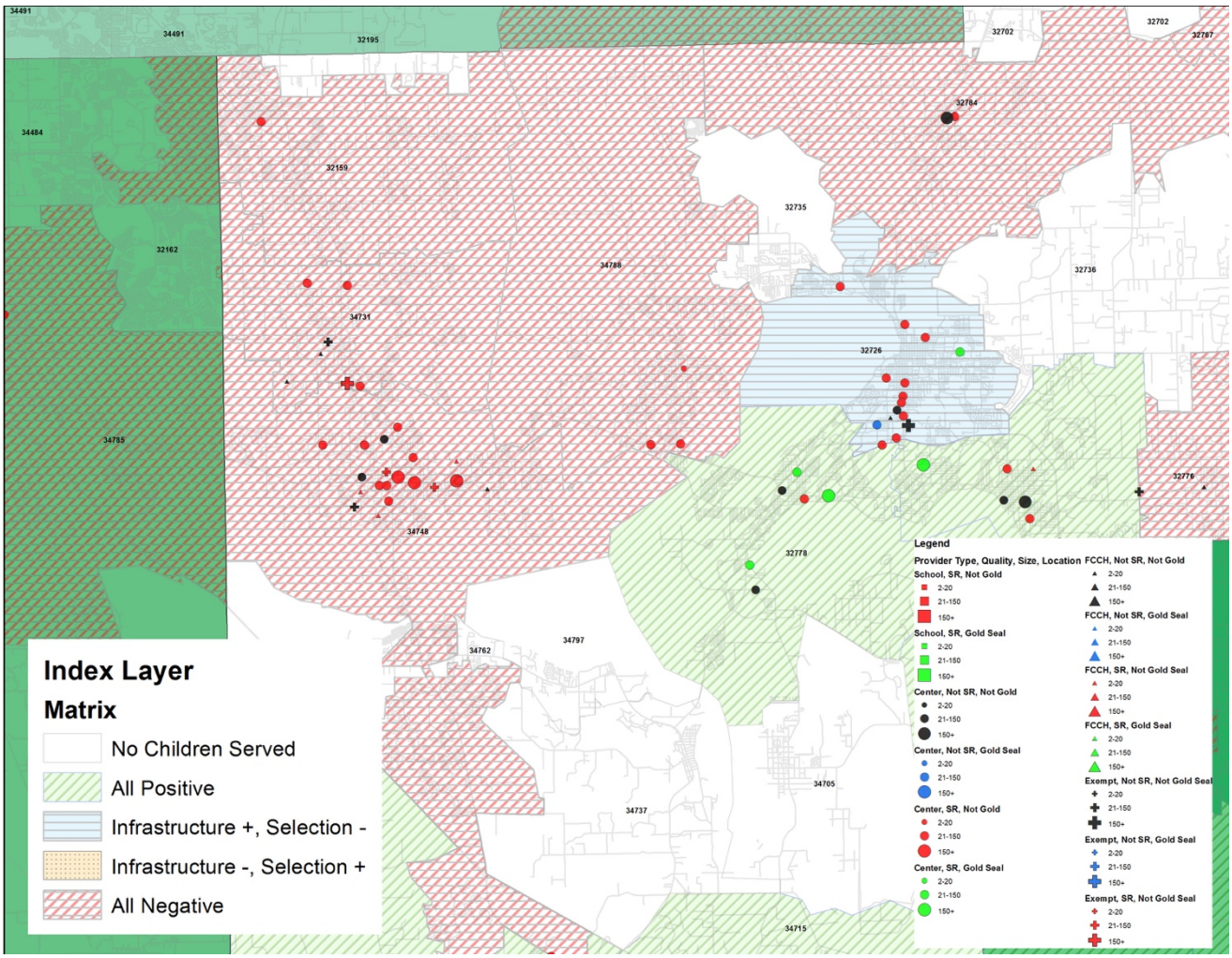
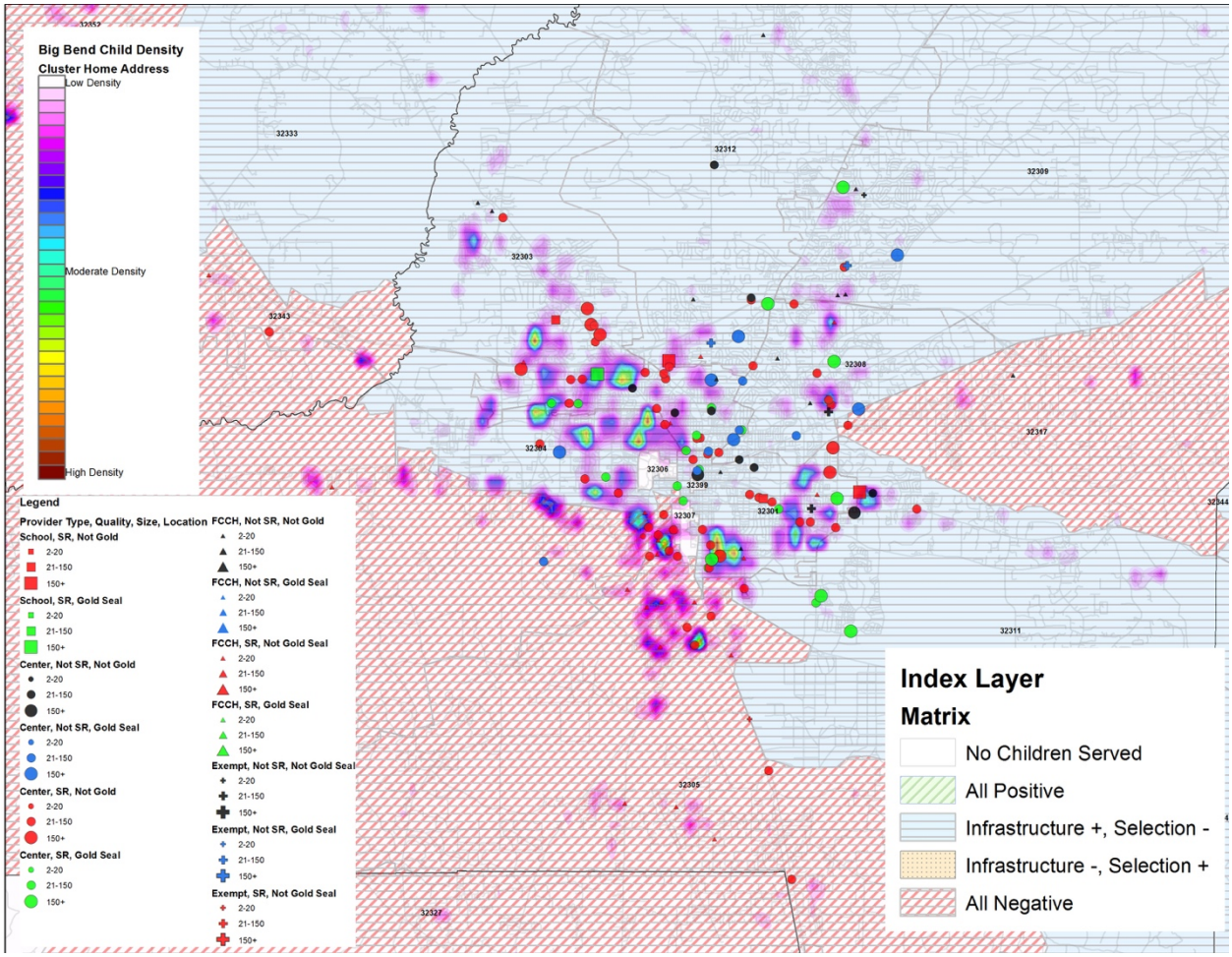


Figure 7: Family Residence Kernel

Bringing it all together

The following map shows all layers at once, that when taken together provide a rich illustration of the extent to which families have reasonable access to high quality early care and education services near to where they live.



Limitations and Lessons Learned

During the course of this project it became apparent that provider capacity information maintained in the OEL database was not robust enough to ensure reliable calculations of the index. In order to bypass this issue, more reliable provider capacity data was obtained from the publicly available DCF data portal. One shortcoming associated with the work to this point involves an incongruence between the dates associated with the OEL subsidy receipt data and the provider information obtained from the FDCF. More specifically, the provider-level capacity information used to create the illustrative analyses in this report were current as of March 2018. However, the most recent subsidy receipt data obtained from OEL was from July 2017. The incongruence resulted in non-overlapping dates of provider information and subsidy use. This resulted in zip code level provider capacity that did not exactly reflect the most accurate provider information for June 2017. As our team moves forward with this partnership, new partnership and data sharing agreements will need to include the Florida Department of Children and Families. Once this partnership is established, we will work to establish a consistent executed process of importing provider information derived from the DCF data system to begin to capture historical, and time specific snapshots of available child care providers.

Recommendations and Future Directions

Integrate Index of Child Care Access with ECENA data portal

While the maps that have been created are tremendously helpful, as reported by the Index Advisory Committee, their utility will be significantly enhanced through integration with the ECENA data portal. Once that integration has taken place the Early Learning Coalitions, the maps will become truly interactive with the capability to showing or hiding various map layers that we have created which will support data informed decision making. Early Learning Coalition and State level program administrators will be able to explore their data in a meaningful way to learn how various elements of their early care and education system is functioning at the local level. Additionally, integration with the ECENA data portal will support deeper analysis of relationships between access to quality child care and other community characteristics such as school district designations, local economic status, and child and family health outcomes.

Revised Monthly Data Extraction and Reporting

Moving forward it is imperative that a parallel data extraction procedure be developed in order to leverage data made available by the OEL and DCF. Data extracted from these two sources at monthly intervals, will offer the collaborating agencies and the research team the opportunity to develop a more reliable and accurate data warehouse. Beyond coordinating simultaneous extractions at monthly intervals, a rigorous curation process based on the lessons learned from the pilot project will ensure that data is processed and cleaned through a series of validations before any indices are calculated. These validations will cover three mains areas. The first check will involve a straightforward screening for missing data across each of the fields in each of the extracted tables. This was an immensely valuable validation measure that led to the discovery of the shortcoming associated with the OEL capacity data. The second check is a macro-level

examination of child-level reporting accuracy. OEL subsidy receipts and the number of children served will be compared to monthly School Readiness reports. This comparison will serve as another reciprocal data integrity validation between agencies, whereby a disparity in the number of subsidy recipients reported between the two agencies will add value to all efforts moving forward. The third check will detail the completeness of the child care provider lists extracted from the OEL and DCF servers. As previously mentioned, this will mitigate the issues encountered in the pilot related to missing provider capacity information in the OEL data extracts.

Develop additional measures pertinent to understanding Child Care Accessibility

Beyond information that is presently available through OEL and DCF, understanding and predicting subsidy demand across geographic areas will require information not only on the number of families and children who receive subsidies but the number of families and children who apply as well. This is critical because at the present moment there does not exist a valid, reliable way to estimate subsidy demand. Rather, being able to contrast the actual number of applicants with the actual number of recipients will offer a much more robust indication of demand that is not prone to measurement error associated with sampling. This robust yet straightforward method for measuring demand is completely in line with the arithmetic indices that led to the development of this project. Furthermore, this sound, defensible way of measuring rather than estimating demand will allow for future robust, predictive models of demand based on indicators of community-based child care markets.

Explore community characteristics to better understand implications for the Index

Once a reliable data collection procedure has been established the next step will be to calculate the indices longitudinally. Tracking how these indices fluctuate over time is the first step in understanding the differential needs of families and local child care markets. Analyzing monthly and seasonal variations in the indices over time will be crucial in bolstering cooperative efforts by stakeholders and the research team to identify intervention strategies. Calculating the indices will require the development of several longitudinal data sets with the acquired monthly data extracts. These longitudinal data sets at the provider, child, and family levels will ultimately be coalesced at appropriate geographic levels that are indicative of natural child care markets within early learning coalitions. This will require a statistical programmer who is both familiar with the data and able to run checks to ensure data integrity throughout a series of merges.

Once a longitudinal data repository for the indices has been established, the next step will be to investigate how these indices co-vary across profiles of community-based child care markets. The goal of these analyses will be to simultaneously discover homogenous subgroups of communities based on a set of relevant economic, child-care specific features and determine how variations in these community profiles are associated with variations in the indices of child care access. Mixture models are an appropriate and intuitive statistical framework for this inquiry. These models will not assume community features have the same effects across communities. Rather, the primary purpose of these analyses will be to identify variations in constellations of community attributes that contribute to differences in how demand, infrastructure, and selection vary across child care markets. More specifically, this approach

does not assume that a particular community-level variable has a constant effect on access to and demand for child care for subsidy recipients across different geographic areas. This holistic approach has been established as a useful approach in market segmentation, as it offers an intuitive interpretation of findings. Furthermore, mixture models will identify communities across geographic areas with similar child care needs based on demographic and economic features. This will be helpful in developing consistent, targeted interventions that meet the needs of communities. Similar to child care markets and the communities in which they operate, the approach to understanding how to best allocate resources must be flexible and evolve to meet the needs children and families across Florida. To this end, the focal point of this project and all analyses will be an ongoing partnership between the research team and the Early Learning Coalitions to develop tailored strategies for improving access to high quality child care.

References

Al Mamun, Md. S & Lownes, Nicholas E. 2011. A Composite Index of Public Transit Accessibility. Transportation, 14 (2): 69-87.

Antle, B. F., Frey, A., Barbee, A., Frey, S., Grisham-Brown, J., & Cox, M. (2008). Child care subsidy and program quality revisited. *Early Education and Development 19*(4), 560-573.

Bambridge, J., Meyers, M.K., & Walfogel, J. (2003). Child care policy reform and the employment of single mothers. *Social Science Quarterly, 84(*4), 771-791.

Baum, C. L. (2002). A dymanic analysis on the effect of child care costs on the work decisions of low-income mothers with infants. *Demography, 39*(1), 139-164.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*, 166-176. doi: 10.1016/j.ecresq.2009.10.004

Blau, D., & Hagy, A. P. (1998). The demand for quality in child care centers. *Review of Economics and Statistics, 84*(3), 483-496.

Blau, D., & Robins, P. (1998). Child care costs and family labor supply. *Review of Economics and Statistics, 70,* 374-381.

Blau, D., & Tekin, E. (2007). The determinants and consequences of child care subsidies for single mothers in the USA. *Journal of Population Economics, 20,* 719-741.

Brooks, F., Reisler, E., Hamilton, C., & Nackerud, L. (2002). Impacts of child care subsidies on family and child well-being. *Early Childhood Research Quarterly 17,* 498-511.

Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early intervention for children fromdifferent social backgrounds. *Early Childhood Research Quarterly, 25,* 140-165.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Report, 112*, 579-620.

Chaudry et al., 2011 Chaudry, A., Pedroza, J. M., Sandstrom, H., Danziger, A., Grosz, M., Scott, M., et al. (2011). *Child care choices of low-income working families*. Retrieved from Urban Institute website: http://www.urban.org/publications/412343.html

Chin, T., & Phillips, M. (2001). Social reproduction and child-rearing practices: Social class, children's agency, and the summer activity gap. *Sociology of Education, 77*(3), 185-210.

Coley, R. L., Li-Grining, C. P., & Chase-Lansdale, P. L. (2006). Low-income families' child care experiences: Meeting the needs of children and families. In N. Cabrera, R. Hutchins, & H. E. Peters (Eds.), From welfare to child care: What happens to young children when mothers exchange welfare for work (pp. 149-170). Mahwah, NJ: Lawrence Erlbaum Associates.

Crosby, D. A., Gennetian, L. A., & Huston, A. C. (2005). Child care assistance policies can affect the use of center-based care for children in low-income families. *Journal of Applied Developmental Science, 9*(2), 86-106.

Dearing, E., McCartney, K. & Taylor, B. A. (2009). Does higher quality early child care promote low-income children's math and reading achievement in middle school? *Child Development, 80,* 1329-1349.

Davis, E.E., Grobe, D., Weber, R.B., (2010). Rural-urban differences in child care subsidy use and employment stability. *Applied Economics Perspectives and Policies*, 32(1), 135- 153.

Davis, E.E., & Weber, R.B. (2001). The dynamics of child care subsidy use by rural families in Oregon.*American Journal of Agricultural Economics, 83*(5), 1293- 1301.

Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning? *School Psychology Review, 36,* 413-432.

Eilers, P. H. C. & Marx, B. D. (1996). Flexible smoothing using B-splines and penalties (with comments and rejoinder). *Statistical Science*, *11*, 89-121.

Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). Regression – Models, Methods and Applications. Springer-Verlag, Berlin.

Fahrmeir, L. & Lang, S. (2001). Bayesian inference for generalized additive models based on Markov random field priors. *Applied Statistics*, *50*(2), 201-220.

Farran, D. (2000). Another decade of intervention for children who are low-income or disabled: What do we know? In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention (Second Edition)* (pp. 510-548). New York: Cambridge University Press.

Forry, N. D., Davis, E. E., & Welti, K. (2013). Ready or not: Associations between participation in subsidized child care arrangements, pre-kindergarten, and Head Start and children's school readiness. *Early Childhood Research Quarterly, 28*, 634-644.

Forry, N. D., & Hofferth, S. (2010). Maintaining work: The influence of child care subsidies on child care related work disruptions. *Journal of Family Issues, 32,* 346-368.

Gilliam, W. S., & Zigler, E. F. (2000). A critical meta-analysis of all evaluations of state-funded preschool from 1977-1998: Implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly, 15,* 441-473.

Gordon, R. A., Kaestner, R., & Korenman, S. (2008). Child care and work absences: Trade-offs by type of care. *Journal of Marriage and the Family, 70,* 239-254.

Gormley, W., & Gayer, T. (2005). Promoting school readiness in Okla- homa: An evaluation of Tulsa's pre-K program, *Journal of Human Resources, 40,* 533–558.

Grobe, D., Weber, R. B., & Davis, E. E. (2008). Why do they leave? Child care subsidy use in Oregon. *Journal of Family and Economic Issues, 29*, 110-127.

Ha, Y. (2009). Stability of child-care subsidy use and earnings of low-income families. *Social Service Review, 83*, 495-525.

Henly et al., 2006 J. Henly, E.O. Ananat, S. Danziger Non-standard work schedules, child care subsidies, and child care arrangements(2006). Retrieved from: http://www.fordschool.umich.edu/research/poverty/pdf/HenDanzAna_Mar2006.pdf

Herbst, C. M. (2008). Who are eligible non-recipients of child care subsidies? . *Children and Youth Services Review, 30*, 1037-1054.

Herbst, C.M. & Tekin, E. (2012). The Geographic Accessibility of Child Care Subsidies and Evidence on the Impact of Subsidy Receipt on Childhood Obesity. Journal of Urban Economics, 71(1), 37-52.

Jones-Branch, J. A., Torquati, J., Raikes, H., & Pope Edwards, C. (2004). Child care subsidy and quality. *Early Education and Development, 15*, 327-342.

Kandala, N.B., Lang, S., Klasen, S. & Fahrmeir, L. (2001). Semiparametric Analysis of the Socio-Demographic Determinants of Undernutrition in Two African Countries. Research in Official Statistics, Vol. 4 No.1:81-100.217.

Kimmel, J. (1998). Child care costs as a barrier to employment for single and married mothers. *Review of Economica and Statistics, 80*, 287-299.

Kneib, T. (2006).Mixed model based inference in structured additive regression. PhD thesis, Dr.Hut-Verlag.

Kneib, T. & Fahrmeir, L. (2006). Structured additive regression for categorical space-time data: a mixed model approach. *Biometrics*, *62,* 109-118.

Knox, V. W., London, A. S., & Scott, E. K.(2003). Welfare reform, work, and child care: The role of informal care in the lives of low-income women and children. Retrieved June 10, 2015, from http://www.mdrc.org/publication/welfare-reform-work-and-child-care

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal*, *104*(5). 409-426.

Lang, S. & Brezger, A. (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics, 13,* 183-212.

Lazar, I. & Darlington, R. B. (1978). *Lasting effects after preschool: A report of the Consortium for Longitudinal Studies.* US Department of Health, Education, and Welfare; Office of Human Development Services; Administration for Children, Youth, and Families.

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D.,… Barbain, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly, 22*, 3-17.

NICHD Early Child Care Research Network. (2000). The relation of child care to cognitive and language development. *Child Development, 71,* 960-980.

NICHD Early Child Care Research Network. (2002). Early child care and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal, 39,* 133-164.

NICHD Early Child Care Research Network & Duncan, G. Y. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74,* 1454–1475.

Nkurunziza, H., Gebhardt, A., & Pilz J. (2011). Geo-additive modeling of malaria in Burundi. *Malaria Journal*, *10:*234.

Pata, M. P., Kneib, T., Cadarso-Suárez, C., Lustres-Pérez, V., & Fernández-Pulpeiro, E. (2012). Categorical structured additive regression for assessing habitat suitability in the spatial distribution of mussel seed abundance. *Environmetrics*, *23,* 75-84.

Peyton, V.; Jacobs, A.; O'Brien, M. & Roy, C. (2001) Reasons for choosing child care: associations with family factors, quality, and satisfaction. *Early Childhood Research Quarterly, 16,* 191-208.

Pungello, E. & Kurtz-Costes, B. (2000) Working Women's Selection of Care for Their Infants: A Prospective Study. *Family Relations, 49*(3), 245-255.

Reynolds, A. J., Magnuson, K. & Ou, S. (2010). Preschool-to-third grade programs and practices: A review of research. *Children and Youth Services Review, 32*(8), 1121-1131

Ryan, R. M., Johnson, A., Rigby, E., Brooks-Gunn, J. (2011). The impact of child care subsidy use on child care quality. *Early Childhood Research Quarterly, 26,* 320-331.

Schweinhart, L. J. (2005). The high/scope perry preschool study through age 40: Summary, conclusions, and frequently asked questions. Ypsilanti, MI: High/Scope Press.

Tekin, E. (2005). Child care subsidy receipt, employment, and child care choice of single mothers. *Economic Letters 89*, 1-6.

Tekin E. (2007). Childcare subsidies, wages, and employment of single mothers. *Journal of Human Resources, 42*, 453-487.

Torquati, J. C., Raikes, H. H., Huddleston-Casas, C. A., Bovaird, J. A., Harris, B. A. (2011). Family income, parent education, and perceived constraints as predictors of observed program quality and parent rated program quality, *Early Childhood Research Quarterly, 26*, 453-464.

Wat, A. (2010). *The case for pre-k in education reform: A summary of program evaluation findings,* Washington, DC: Pre-K Now, Pew Center on the States. Retrieved from http://www.preknow.com/documents/thecaseforprek_april2010.pdf

Weber, R.B., & Grobe, D. (2011). *Oregon subsidy policy impact research project: Parent survey.* Retrieved on June 11, 2014 from: http://health.oregonstate.edu/sites/default/files/sbhs/pdf/oregon-subsidy-policy-impact-research-project-parent-survey-report-8-22-2011-revised.pdf

Appendices:

Appendix A: Data Codebook

Appendix B: Alachua County

Appendix C: Big Bend Early Learning Coalition

Appendix D: Lake County Early Learning Coalition

Appendix E: Pinellas County Early Learning Coalition

Appendix F: Southwest Florida Early Learning Coalition

Appendix G: Index Table

Appendix A: Data Codebook

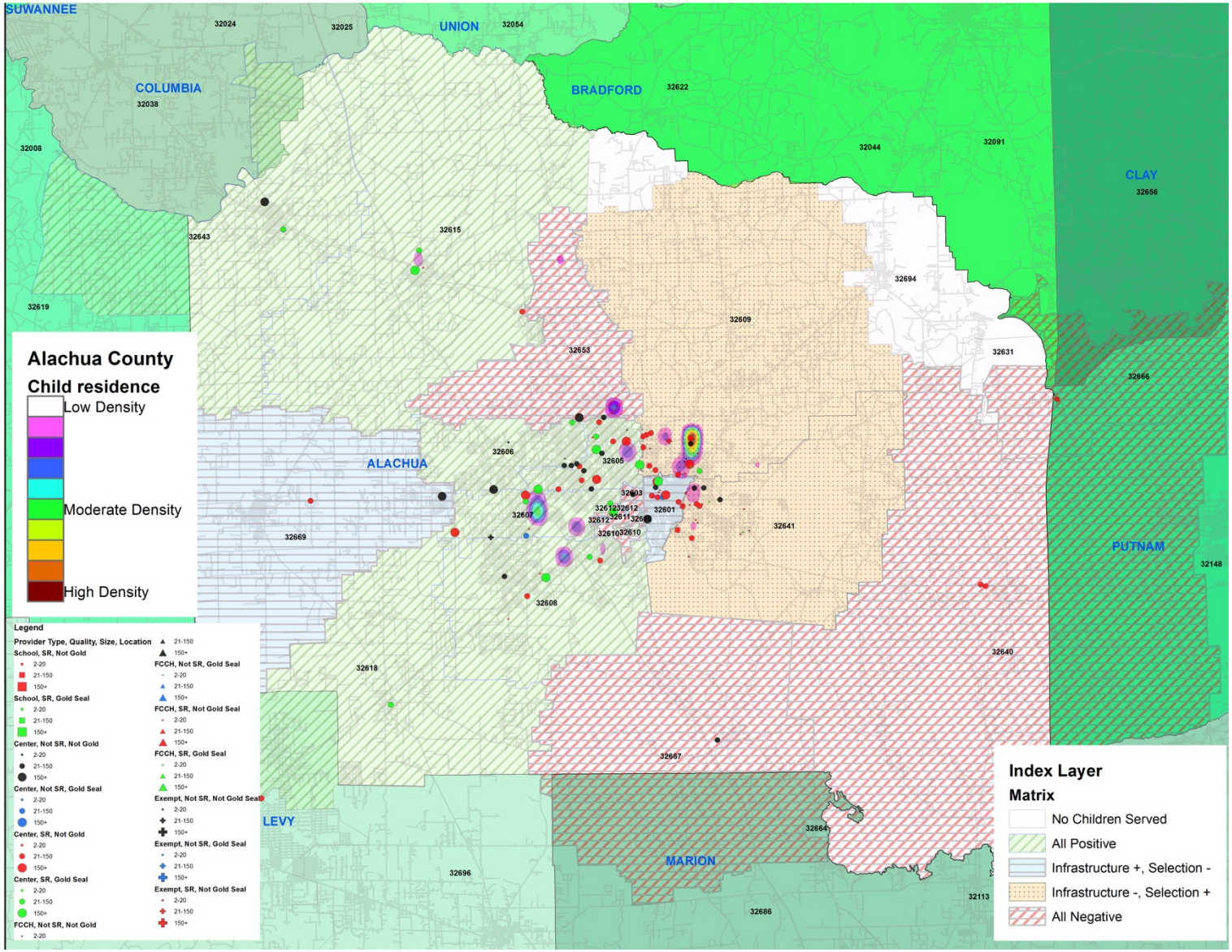
Child File Codes

Child	
Field	**Description**
COALITION_NAME	Name of the Early Learning Coalition.
PRRP_D_START	First day of the month when service was provided. Covers the entire month.
COALITION_ID	Database number. School Readiness and VPK data are on 35 separate databases.
CHLD_C_SSN	Child SSN or system-generated ID. Children are uniquely identified by the combination of COALITION_ID and CHLD_C_SSN.
CHLD_F_SSN	Flag to indicate if the Child's ID is an SSN (Y) or not (N).
CHLD_D_DOB	Child Date of Birth.
PRNT_C_SSN	Parent SSN or system-generated ID. Parents are uniquely identified by the combination of COALITION_ID and PRNT_C_SSN.
PRNT_F_SSN	Flag to indicate if the Parent's ID is an SSN (Y) or not (N).
PRNT_L_RES_STRT	Parent's Residence Street Address.
PRNT_L_RES_STRT2	Parent's Residence Street Address - optional second line.
PRNT_L_RES_CITY	Parent's Residence City.
PRNT_L_RES_ST	Parent's Residence State.
PRNT_L_RES_ZIP	Parent's Residence Zip Code.
PRRD_C_CRLV_ABV	Child's Care Level. See 'Care Level Codes' Tab.
CHLD_F_HEADSTART	Flag to indicate the child is attending Head Start (Y) or not (N or Blank).
PROV_C	Provider ID (either FEID or SSN for Family Child Care Homes).
PROV_C_EXT	Provider Extension - Used to uniquely identify providers with multiple locations under the same FEIN. Providers are uniquely identified by the combination of COALITION_ID, PROV_C, and PROV_C_EXT.
PROV_C_TP	Provider Type at the time of the payment calculations. Note that a provider could change types over time. This is the provider type at the time the service was provided.
GOLD_SEAL	Y' means the Provider has a Gold Seal Provider Type, 'N' means not Gold Seal.
PROV_GROUP	Provider Types grouped into broader categories.
FNDR_C	Funder. 1= School Readiness, VPK = Voluntary Prekindergarten.
FDCN_C	Billing Group - See the 'School Readiness Billing Groups' Tab and the 'VPK Billing Groups' Tab. Children were reported for receiving School Readiness or VPK payments at any time between 7/1/2015 and 6/30/2017.
ELIG_C	Eligibility is a subdivision of the Billing Group. See the 'School Readiness Billing Groups' Tab and the 'VPK Billing Groups' Tab for details.

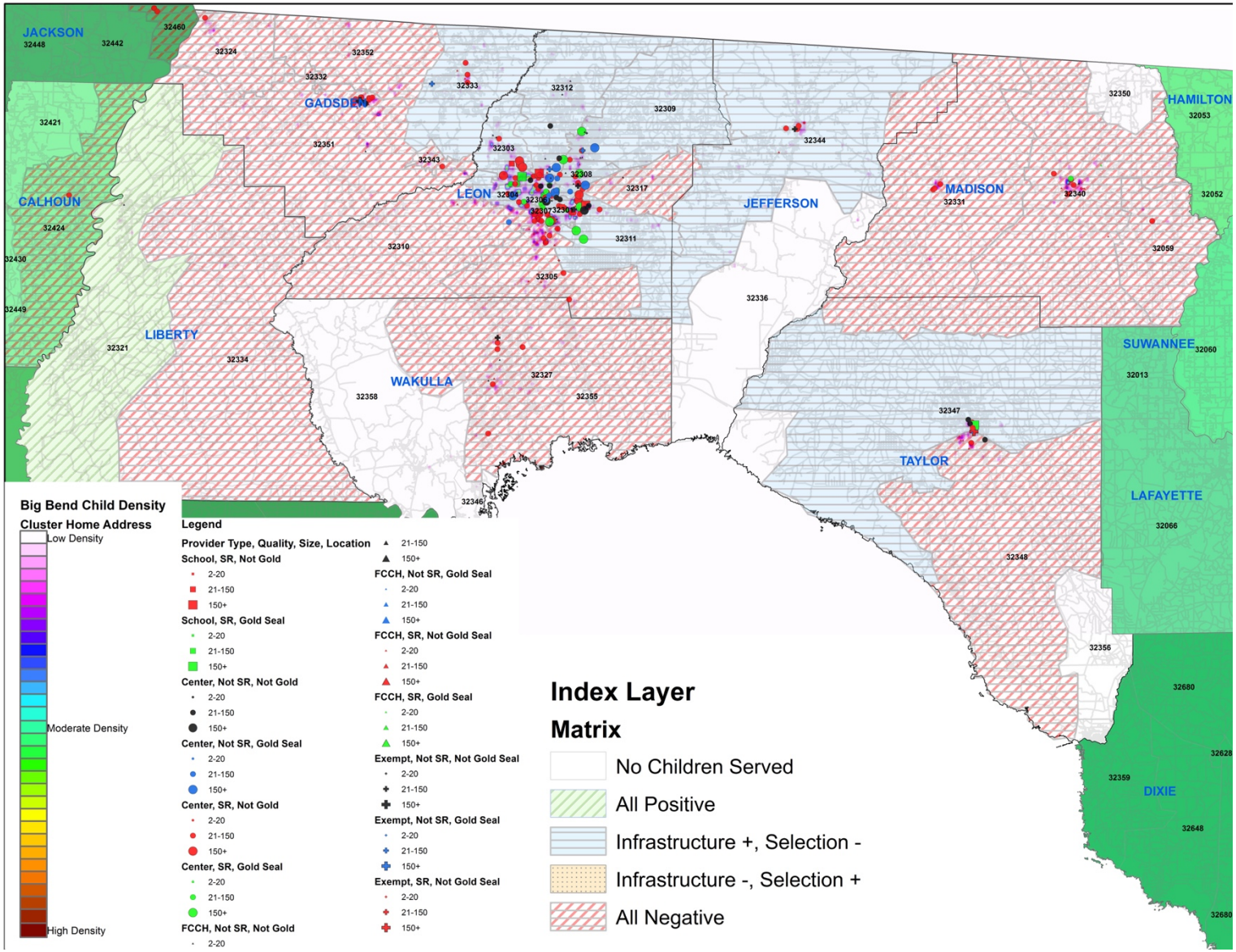
Provider file codes

Provider	
Field	**Description**
COALITION_NAME	Name of the Early Learning Coalition.
COALITION_ID	Database number.
PROV_C	Provider ID (either FEID or SSN for Family Child Care Homes).
PROV_C_EXT	Provider Extension - Used to uniquely identify providers with multiple locations under the same FEIN.
PROV_C_LEGAL_ID	Legal ID Number. Field is not required. If absent, use PROV_C_DCF_ID.
PROV_C_DCF_ID	DCF ID Number.
PROV_N	Provider Name - Payment purposes.
PROV_N_CCRR	Provider Name - Public-Facing.
PROV_L_STRT	Provider's Physical Street Address.
PROV_L_STRT2	Provider's Physical Street Address - optional second line.
PROV_L_CITY	Providers Physical City.
PROV_L_ST	Provider's Physical State.
PROV_L_ZIP	Providers Physical Zip Code.
PROV_C_TP	Provider's Current Provider Type - See Standard Codes. This is the most recent Provider Type.
GOLD_SEAL	Y' means the Provider currently has a Gold Seal Provider Type, 'N' means not Gold Seal.
PROV_GROUP	Provider Types grouped into broader categories.
PREN_Q_CAP_REAL	Actual Capacity for all Age Levels.
PREN_Q_CAP_LIC	Licensed Capacity for all Age Levels.

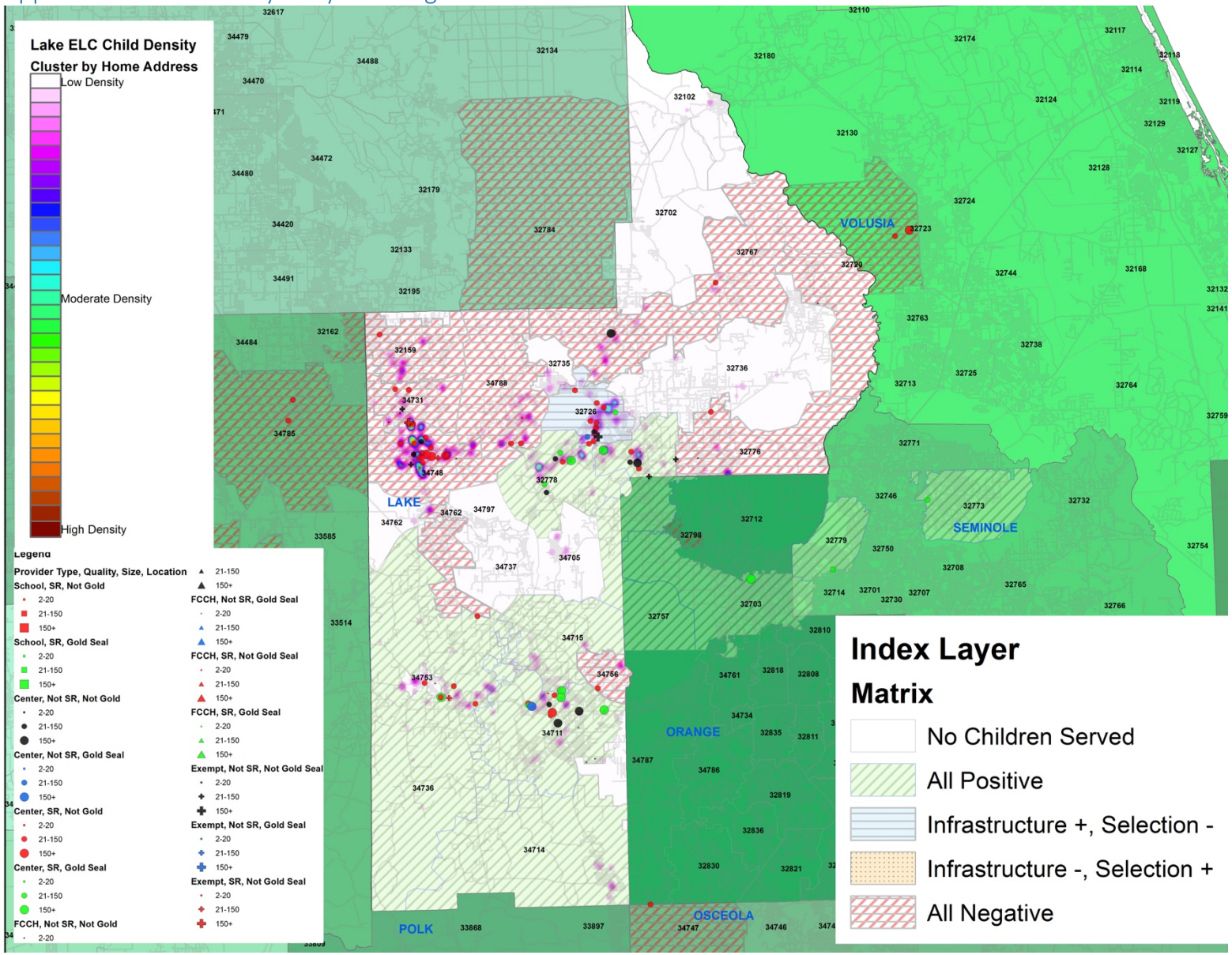
Appendix B: Alachua County



Appendix C: Big Bend Early Learning Coalition

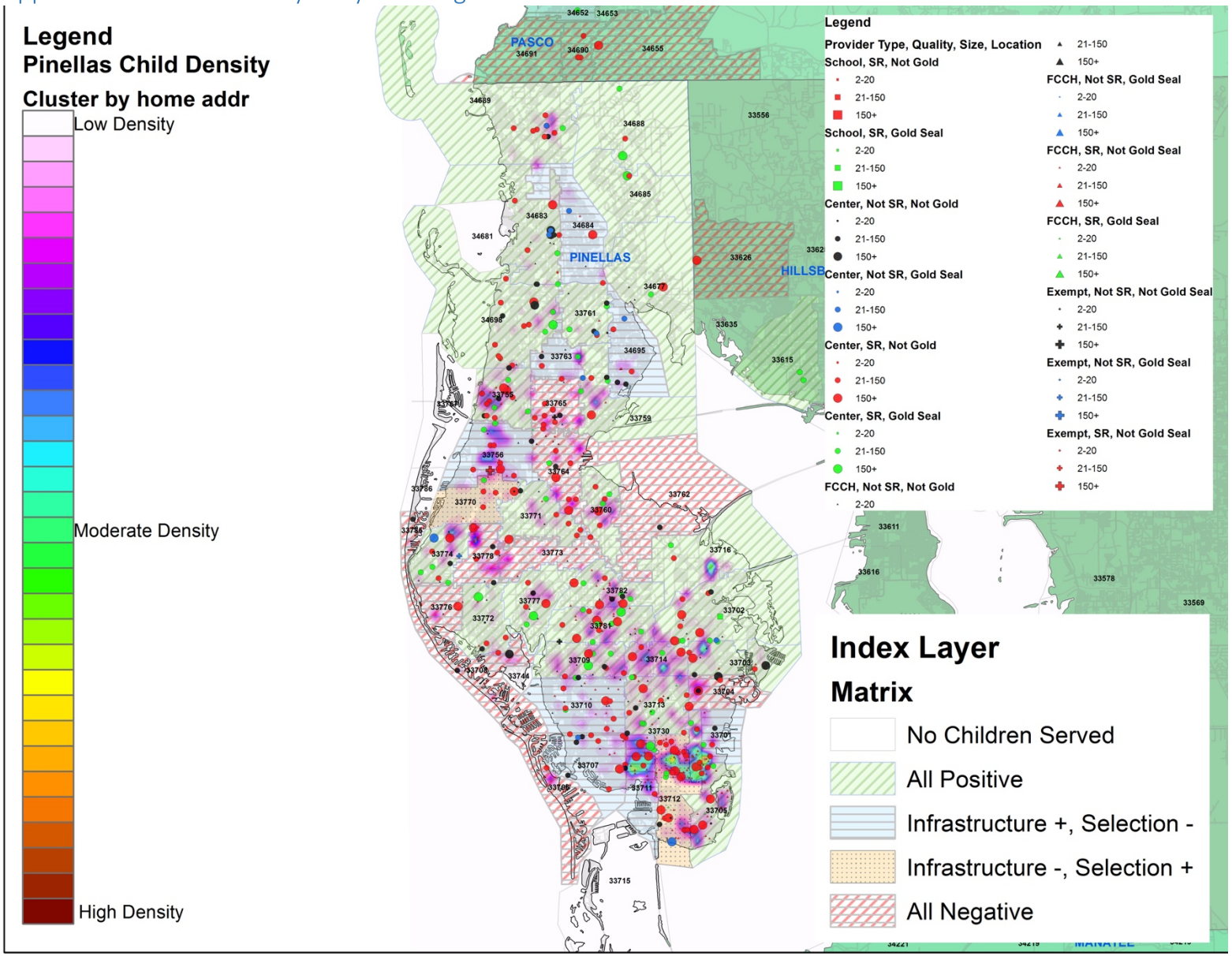


Appendix D: Lake County Early Learning Coalition

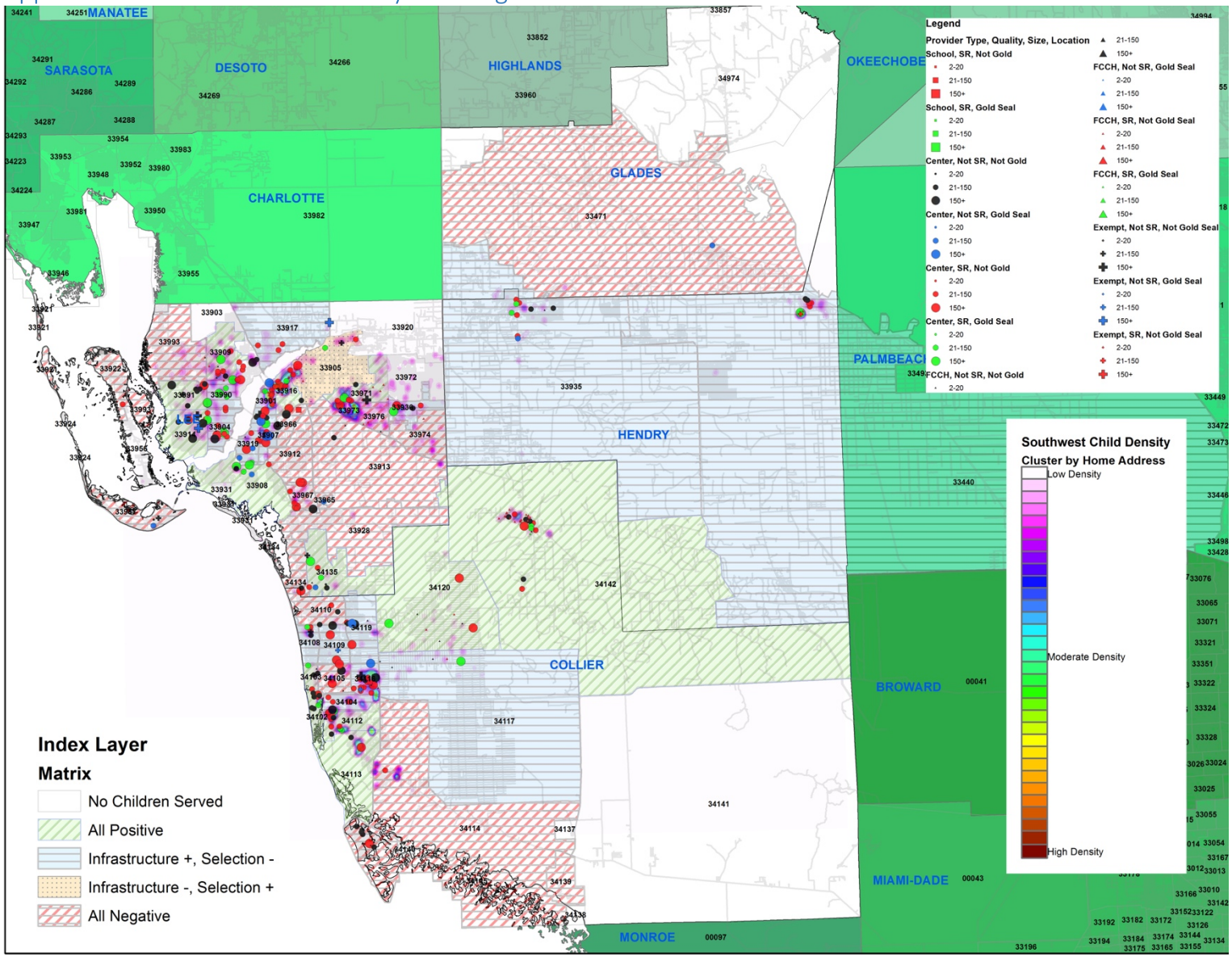


Prepared for the Florida Office of Early Learning by The University of Florida Childhood Needs Assessment Partnership – OEL Contract # SR660 – July, 2018

Appendix E: Pinellas County Early Learning Coalition



Appendix F: Southwest Florida Early Learning Coalition



Appendix G: Index Table

Zip Code	Gold Seal Capacity	Enrolled Gold Seal	Not Gold Capacity	Enrolled Not Gold	Selection	Infrastructure	Quadrant
32301	1033	79	1732	392	-0.149851662	562	2
32303	991	152	2473	467	-0.035459042	372	2
32304	411	56	274	161	-0.4513382	194	2
32308	1002	27	1144	182	-0.132144801	793	2
32309	721	2	310	16	-0.048838978	703	2
32310	76	4	722	329	-0.403047091	-257	4
32311	616	42	563	34	0.007791054	540	1
32312	756	18	1008	33	-0.008928571	705	2
32321	242	3	10	0	0.012396694	239	1
32333	95	0	245	82	-0.334693878	13	2
32340	60	19	301	143	-0.15841639	-102	4
32344	97	5	251	114	-0.402636875	-22	4
32347	249	15	439	56	-0.067321679	178	2
32348	20	0	172	61	-0.354651163	-41	4
32601	423	88	776	214	-0.067735371	121	2
32605	1183	175	1377	108	0.069497622	900	1
32606	241	24	384	0	0.099585062	217	1
32607	315	41	545	80	-0.016630261	194	2
32608	450	109	419	35	0.158690003	306	1
32609	141	69	1068	252	0.253406646	-180	3
32610	125	0	0	0	0	125	4
32611	88	0	0	0	0	88	4
32612	189	4	0	0	0.021164021	185	1
32615	283	54	130	19	0.044658875	210	1
32618	67	19	89	2	0.261110179	46	1
32641	22	21	830	248	0.655750274	-247	3
32643	72	18	224	7	0.21875	47	1
32669	202	0	249	11	-0.044176707	191	2
32703	402	4	0	0	0.009950249	398	1
32726	217	35	1144	220	-0.03101737	-38	4
32757	230	87	981	54	0.323214998	89	1
32773	130	2	0	0	0.015384615	128	1
32778	453	113	291	53	0.067317539	287	1
32779	99	1	0	0	0.01010101	98	1
32818	NA	4	0	0	NA	NA	NA
33440	294	35	615	96	-0.037049942	163	2
33471	185	0	0	0	0	185	4
33615	212	4	0	0	0.018867925	208	1
33701	138	2	399	63	-0.143401983	73	2
33702	564	71	1258	63	0.075807034	430	1
33703	134	61	806	18	0.432891374	55	1
33705	99	32	1925	100	0.271284271	-33	3
33707	186	0	960	161	-0.167708333	25	2
33709	839	147	706	102	0.30233259	80	1

Acknowledgements

Alachua County Early Learning Coalition
Big Bend Region Early Learning Coalition
Lake County Early Learning Coalition
Pinellas County Early Learning Coalition
Southwest Florida Region Early Learning Coalition
UF Research Computing
UF GeoPlan Center
UF Division of Sponsored Research

OFFICE OF
Early Learning
LEARN EARLY. LEARN FOR LIFE.

UF | Anita Zucker Center *for* Excellence *in* Early Childhood Studies
UNIVERSITY *of* FLORIDA

UF | Family Data Center
UNIVERSITY *of* FLORIDA

Prepared for the Florida Office of Early Learning by The University of Florida Childhood Needs Assessment Partnership – OEL Contract # SR660 – July, 2018