

# National Survey of Early Care and Education:

## Summary of Data Collection and Sampling Methodology



NSECE RESEARCH BRIEF

OPRE REPORT #2013-46 | OCTOBER 2013



National Survey of **Early Care & Education**

**OPRE**

**Disclaimer**

The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

# National Survey of Early Care and Education: Summary Data Collection and Sampling Methodology

## NSECE RESEARCH BRIEF

OPRE REPORT 46 | OCTOBER 2013

Submitted to:

Ivelisse Martinez-Beck, Ph.D., Project Officer  
Ann Rivera, Ph.D., Research Analyst

Office of Planning, Research and Evaluation  
Administration for Children and Families  
U.S. Department of Health and Human Services

Contract Number: HHHSP23320095647WC

Project Director:

A. Rupa Datta  
NORC at the University of Chicago  
55 E Monroe Street  
Chicago, Illinois, 60603

**National Survey of Early Care and Education Project Team. (2013).** National Survey of Early Care and Education: Summary Data Collection and Sampling Methodology. OPRE Report #2013-46, Washington DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

### Authors:

*Marietta Bowman*  
*Jill Connelly*  
*Rupa Datta*  
*Lekha Venkataraman*  
*Micah Sjoblom*  
*Marilyn Brooks*  
*Kanru Xia*  
*Joshua Borton*  
*Hee-Choon Shin*  
*Kirk Wolter*

This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at <http://www.acf.hhs.gov/programs/opre/index.html>.



Table of Contents

A. Study Overview ..... 1

B. Sample Design ..... 1

C. Component Surveys ..... 5

D. Additional Distinctive Features..... 8

E. Project Timeline..... 10

## A. STUDY OVERVIEW

The primary purpose of the National Survey of Early Care and Education (NSECE) was to provide a comprehensive snapshot of both the availability and utilization of early care and education in the United States. The main objectives of the study included:

- Providing the first national portrait of the availability of early care and education for the full spectrum of care providers, including households and providers from all 50 states and the District of Columbia.
- Identifying early care and education and school-age care (ECE/SA) needs and preferences among households in the United States with children under age 13 as they pertain to supporting both the employment of parents and the development of children.
- Capturing data on all forms of non-parental care for all children in a household.
- Providing the perspectives of both families and providers on the services offered in a system where children are often in multiple arrangements and providers receive funding from multiple sources.
- Linking the data set collected with policy-relevant data.
- Increasing the understanding of the care received by low-income children and how that varies across communities.

The NSECE is a set of four integrated, nationally representative surveys conducted in 2012. These were surveys of 1) households with children under 13, 2) home-based providers 3) center-based providers, and 4) the center-based provider workforce. Together they characterize the supply of and demand for early care and education in America and permit better understanding of how well families' needs and preferences coordinate with providers' offerings and constraints. The study is funded by the Office of Planning, Research and Evaluation (OPRE) in the Administration for Children and Families (ACF), US Department of Health and Human Services. The project team is led by NORC at the University of Chicago, with partners Chapin Hall at the University of Chicago and Child Trends, as well as other collaborating individuals and organizations.

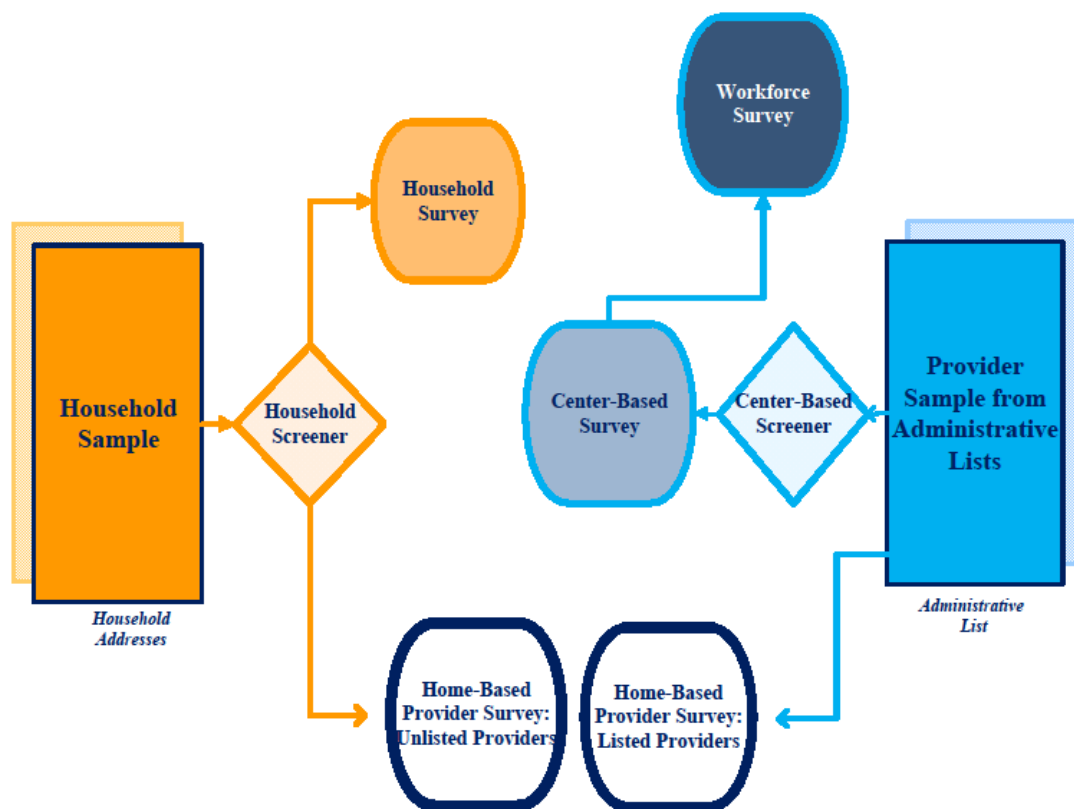
## B. SAMPLE DESIGN

Figure 1 provides an overall schematic of the NSECE sample types and questionnaires. The NSECE is a coordinated set of four nationally representative surveys pertaining to the supply of and demand for early care and education in the United States, including the individuals working directly with children. There are two primary sources of sample for these four surveys. A household sample was constructed as a hybrid between an address-based sample of housing units selected from the Delivery Sequence File (DSF) maintained by the US Postal Service and a freshly listed sample of housing units in a small number of locations where the DSF lacked adequate coverage to support a high-quality sample. Using a household screener, eligible households were identified for the household questionnaire and for the home-based provider questionnaire from this household sample.

In order to draw a nationally representative sample of the supply of early care and education, the project constructed a sampling frame of "listed" providers from administrative lists. This frame was built through compiling and geo-coding all available state-level and national lists of providers of early care and education collected from various agencies in all 50 states and the District of Columbia. These lists of providers included licensing, regulation, and license-exempt lists, as well as lists of providers in specific programs such as offering Head Start or public pre-

kindergarten. Three different surveys used this sample source. Center-based providers of early care and education to children not yet in kindergarten were selected through a center-based screener for the center-based provider questionnaire. From the center-based providers who completed a center-based provider interview, respondents were selected for the workforce questionnaire. Also from the administrative lists, home-based providers were selected for the home-based provider survey. Note that the home-based provider survey includes sample from both sample sources: the household (for unlisted providers) and the administrative lists (for listed providers).

**Figure 1 NSECE Sample Types and Questionnaires**



The NSECE sample design is a multistage probability design. In the first stage, 219 primary sampling units (PSUs) were selected across all 50 states and DC. PSUs were allocated to states by size based on the population of children under age 18 within each state. In the second stage, secondary sampling units (SSUs) were selected for the household sample. Because the experiences of low-income families are of special interest in public policy addressing early care and education/school-age (ECE/SA), the NSECE sample design included a low-income oversample. SSUs were selected disproportionately from areas in which at least 40 percent of households had income below 250 percent of federal poverty guidelines. Altogether, 755 SSUs were selected, with 537 SSUs in these lower-income areas and 218 in areas with lower densities of low-income households.

The NSECE sample design introduces the concept of a provider cluster to generate nationally representative estimates while capturing the very local nature of how families seek and use

ECE/SA, how providers seek and serve children, and how these things together affect the context in which ECE utilization occurs. The map below depicts a hypothetical cluster using a location in Dallas, Texas. The SSU is the central yellow area, which represents the cluster's core of households, while the gray shaded areas depict the remainder of the cluster. Households in the yellow core (generally one or a small number of adjacent census tracts) were sampled for inclusion in the household and home-based provider questionnaires. The provider sampling frame built from administrative lists was used to sample listed providers, including center-based programs and home-based ECE providers, from throughout the gray and yellow portions of the cluster, approximating the locations from which the centrally located households might seek ECE services. The gray portion comprises all census tracts that overlap within a circle of two miles centered at the population centroid of the yellow core. The use of the provider cluster allows us to document the interaction of the supply of and the demand for early care and education where it occurs—in local communities—while simultaneously capturing data that efficiently construct national estimates.

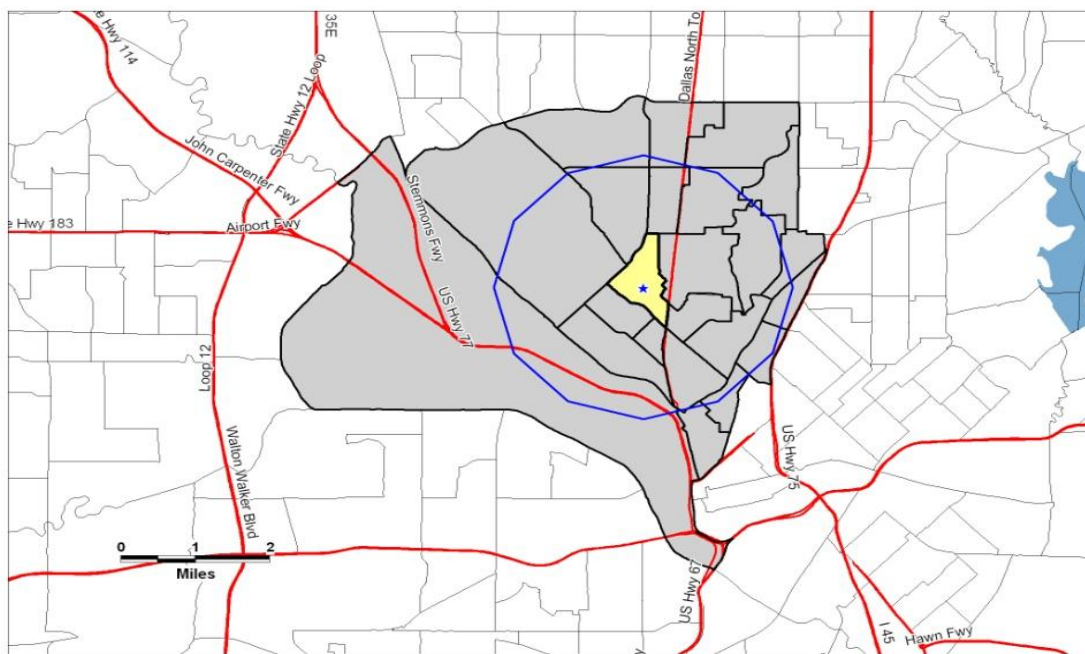
OPRE made available to the states the opportunity to supplement their NSECE samples for the purpose of increasing state-specific sample sizes and analytic power. The states of New York and Illinois both exercised this option to supplement. The 219 PSUs and 755 SSUs in the final sample reflect an expansion of the number of PSUs by two and the number of SSUs by 14 relative to what would have been allocated in the absence of supplementation.

Additional information about the design of the NSECE sample is available in the Revised Sampling Report and Addendum prepared during the design phase of this study:

[http://www.acf.hhs.gov/sites/default/files/opre/revised\\_sampling.pdf](http://www.acf.hhs.gov/sites/default/files/opre/revised_sampling.pdf).

**Figure 2 Hypothetical Provider Cluster**

## Dallas County 0006.01





*Provider sampling.* The NSECE team built a comprehensive file of early care and education (ECE) programs serving children under age 13 in the 50 states and the District of Columbia. In each state, the child care licensing unit, division, or department was contacted and asked to provide all available lists of licensed, registered, or otherwise compiled child care providers. To supplement state lists and cover common exemptions, we also collected the following national lists:

- Department of Defense child care
- General Services Administration child care on federal property
- National Association for the Education of Young Children accredited programs
- Boys and Girls Clubs of America
- YMCA child care and afterschool programs
- Office of Head Start national list of programs

We also included a proprietary list of all elementary schools in the nation offering at least one grade K through 8 and any early childhood program operated by a public school district. These were included as potential providers of early care and education (for example, pre-kindergarten or school-age care), although regular elementary school itself was not sufficient to qualify for the Center-Based Provider survey.

Child care licensing list collection occurred primarily from February to June 2011. We obtained child care licensing lists from all 50 states and Washington, DC. We also documented all list types and exemptions in each state. The common lists obtained for states were Home-based family or group care and center care. Pre-K lists were collected primarily from April to August 2011. We collected 45 pre-K lists, including three states for which a separate pre-K designation was available in child care licensing lists. Six states did not have a pre-K list available (Arizona, Arkansas, Idaho, Mississippi, Montana, and Wyoming). In most cases, a list did not exist because there was no coordinated state-level funding program for pre-K. With the exception of Montana, where programs serving children over age three and primarily educational in purpose were exempt from licensing requirements, pre-K programs operating outside of public schools should have been included in licensing lists though not separately designated. Mississippi also exempted from licensing Head Start programs operating in public schools.

Afterschool list collection occurred primarily from May to September 2011, though it commenced earlier in the year and then was concentrated in July and August. We obtained 31 afterschool lists, including three states for which a separate license type or code was included in child care licensing lists. There are no states that exempt school-age care for which we were unable to obtain a list.

From the comprehensive file, we extracted a sampling frame consisting of all unique addresses housing at least one provider on the sampling frame within the pool of selected provider clusters within each PSU. The ultimate sampling unit for the center-based providers was the organization operating an ECE program at an address. For locations/addresses with multiple programs, we rostered the programs and a single organization operating one or more eligible programs was randomly selected for interview.

*Household Sampling.* We used a delivery sequence file (DSF) maintained by the United States Postal Service (USPS) as the sampling frame for housing units (HUs) at the third stage of sampling. The DSF is known to be incomplete in some areas of the country, especially in some



rural areas. NORC's sampling strategy was to employ the DSF in areas in which it was thought to be reasonably complete and to employ a List & Go approach in the 25 SSUs in which the DSF was thought to be deficient because its count of households was unacceptably smaller than the count of households for that location in the American Community Survey.

## C. COMPONENT SURVEYS

Below, we describe each of the four surveys briefly.

The **Household Survey** was conducted with a parent or guardian of a child or children under age 13 in households with at least one member child under age 13. Eligible respondents were identified through the **Household Screener** based only on the presence of an age-eligible child. Screening was completed by mail, by phone, and in person. The household questionnaire data include 11,629 interviews with adults in households with children under age 13. All of these interviews were conducted by an interviewer, primarily in person with a small fraction by telephone. A total of 65,712 screening interviews were completed, for a weighted screener completion rate of 91.1 percent. From these, 11,629 eligible households completed a Household interview, yielding a weighted interview completion rate of 67.1 percent. The overall weighted response rate is 62.2 percent<sup>1</sup>.

This survey documents the nation's demand for early care and education services. Key questionnaire topics include details on usage of non-parental care, expenditures on non-parental care, parental search behavior for early care and education, and the balance of parental employment with child care needs and availability. These data will help to answer such research questions as 1) Who is caring for America's children when they are not with their parents and do families with different demographic characteristics have different preferences or different patterns of usage? 2) How do families search for care and how does this vary by age of children, characteristics of parents, location, and availability of licensed slots per population? 3) How and how much do families pay for care? and 4) How many families of different characteristics receive public financial support for ECE, and how does this vary by age of child and type of care utilized?

Distinctive features of the household questionnaire include collection of data on all children under age 13 (not just a focal child) and collection of child care payment data at the child-provider pair level rather than in aggregate. The NSECE data offer larger samples of low-income children than do many other sources. The NSECE data are also valuable for more intensively investigating some of the patterns observed in other data. For example, the NSECE data expand the possibilities for understanding how parents coordinate work and school schedules with early care and education usage, and the extent to which different types of care solve or present schedule coordination problems. Data from multiple children, details of parental searches for care, and innovative approaches for determining likely participation in government programs (such as CCDF, Head Start, or public pre-K) are all innovations in the household questionnaire instrument.

The **Home-Based Provider Survey** was conducted with individuals who provide care in a home-based setting for children under age 13 who are not their own. Two sample sources contributed providers. One, home-based providers who appeared in provider sampling frames constructed from state and national were sampled for this survey. We designated these

---

<sup>1</sup> Overall response rates are calculated using the definition for response rate #2 according to guidelines published by the American Association of Public Opinion Research.

providers as 'listed.' Listed providers were primarily licensed or regulated family day care providers, but also included other formally listed home-based providers such as license-exempt providers or providers participating in Early Head Start. Alternatively, households could have been identified as eligible based on their responses to the household screener (specifically, that an adult in the household regularly cared for children not his or her own at least five hours per week in a home-based setting). These providers did not appear in the provider sampling frame and are designated as 'unlisted.' Data collection was conducted by Internet, in person, and by telephone with field interviewers.

The NSECE data include a combined total of 5,986 for Listed and Unlisted Home-Based provider interviews. For Listed Home-Based providers, eligibility was confirmed for a total of 5,752 home-based providers, for a weighted screener completion rate of 86.51 percent. From these, 3,934 eligible Listed Home-Based providers completed a Home-Based provider interview, yielding a weighted interview completion rate of 93.3 percent. The overall weighted response rate is 80.7 percent. For Unlisted Home-Based providers, a total of 65,712 screening interviews were completed, for a weighted screener completion rate of 91.1 percent. From these, 2,052 eligible Unlisted Home-Based providers completed an Unlisted Home-Based provider interview, yielding a weighted interview completion rate of 66.4 percent. The overall weighted response rate for unlisted providers is 67.5 percent.

Key questionnaire topics in the home-based provider questionnaire include enrollment and the characteristics of the children served, rates charged for care, participation in government programs, household composition, qualifications for and attitudes toward early childhood education, use of curricula and activities conducted with children (varied to be appropriate for younger children and school-age children). Portions of the home-based provider questionnaire will contribute to analyses of the ECE workforce and mirror the content of the workforce questionnaire administered to classroom-assigned instructional staff at center-based providers. Other portions of this questionnaire closely mimic the center-based provider questionnaire (see below), so that enrollment, program participation, provider charges for care, attitudes, orientation and activities data can be accurately compared across all different categories of provider. These data will answer such questions as 1) What kind of early care and education is available across communities throughout the country? 2) How well does the available supply of early care and education support parents' employment? 3) How do different types of providers vary in their characteristics of care and affordability? and 4) Who are the individuals working in early care and education? What are their experiences in terms of employment characteristics, classroom activities, and professional development? What are their attitudes, orientations, and stress and depression levels?

The inclusion of providers identified through the household screener offers unprecedented nationally representative data on family, friend and neighbor providers and is one of the pioneering aspects of the NSECE. We will learn about both paid and unpaid care, including how they differ in their characteristics and their availability to families.

The **Center-Based Provider Survey** was conducted with directors of ECE programs that provide care to children not yet in kindergarten who were identified from the provider sampling frame built from state or national administrative lists such as state licensing lists, Head Start program records, or pre-K rolls. These providers included regulated, licensed, and other private providers as well. The center-based provider data come from about 8,265 center-based provider questionnaire interviews. The center-based provider questionnaire was preceded by a **Center-Based Provider Screener** that confirmed and updated information from the provider

sampling frame, determined the availability of school-age after-school care at elementary schools, and determined eligibility for the center-based provider questionnaire. Topics covered by this instrument include enrollment and characteristics of children served, staffing, prices charged, schedules of service, participation in government programs, and staff compensation and professional development policies. The questionnaire also includes the selection of a representative classroom about which more detailed staffing, compensation, and curriculum information are collected. These data constitute a nationally representative sample of ECE classrooms. This instrument was administered using Web, in-person, and telephone interviewing. A total of 15,805 screening interviews were completed, for a weighted screener completion rate of 94.3 percent. From these, 8,265 eligible Center-Based Providers completed a Center-Based interview, yielding a weighted interview completion rate of 78.2 percent. The overall weighted response rate is 73.7 percent.

Although no observational data are collected on the care provided, the questionnaire includes a variety of measures at both the program and individual staff levels that have been found in the literature to predict observed quality of care. No nationally representative data on the supply of early care and education have been available since the 1990 Profile of Child Care Settings. The NSECE data update those data in many ways, but also reflect many contemporary issues, including the blending of public funding sources (sometimes with private funds), the provision of public prekindergarten in school-based and community-based settings, and targeted accommodations such as comprehensive services and services for English-language learners and their families. Selected segments of the center-based provider questionnaire were designed in parallel with the home-based provider questionnaire, so that comparable data would exist for more formal home-based providers as well as for centers. The center-based provider questionnaire data will answer such questions as 1) What kind of early care and education is available across communities throughout the country? 2) How well does the available supply of early care and education support parents' employment? 3) How do different types of providers vary in their characteristics of care and affordability? and 4) How many and what types of providers participate in quality improvement efforts such as staff quality ratings and professional development?

The **Workforce Provider Survey** sample comprised one classroom-assigned instructional staff person from each center-based provider who completed a center-based provider interview. Workforce respondents were drawn from the center-based provider questionnaire data, in which all staff members had been enumerated from a randomly selected classroom. The questionnaire closely mirrors portions of the home-based provider questionnaire, so that the two data sources together can paint a rich portrait of the paid ECE workforce, including center-based and home-based paid providers. (Individuals who were not paid will be profiled as described in the family, friends, and neighbor section above.) Topics include information about the work setting (activities in the classroom, interactions with parents and other staff, availability of professional development and other supports), roles and responsibilities (lead teacher, teacher, assistant teacher, aide), compensation (wages and benefits), and perceived leadership and morale, as well as personal information about qualifications, attitudes toward ECE, and stress, depression, and demographic information. Altogether, 5,556 interviews were completed with workforce respondents using Web, paper-and-pencil, telephone, and in-person modes. A total of 7,230 center-based provider questionnaires were completed with adequate data to sample a workforce respondent, for a weighted screener completion rate of 88.1 percent. From these, 5,556 eligible workforce employees completed a Workforce interview, yielding a weighted interview completion rate of 80.7 percent. The overall weighted response rate is 71.2 percent.

The workforce questionnaire is the simplest of the NSECE survey instruments, but these data represent a signal contribution of the study. No nationally representative survey of paid ECE workers has ever been done, yet many of the pressing ECE policy questions require better understanding of home-based care providers' qualifications, access to professional development, motivation for working in the field, and nature of participation in the labor market as well as the extent to which paid home-based providers are more generally available for expanding and improving overall ECE supply. Some of the workforce questionnaire data will allow tabulation by provider program characteristics (such as enrollment size, type of care, geographic location, for-profit/not-for profit status, and participation in government programs) of factors that have been found in the literature to predict observed quality. These factors include staff qualifications and compensation, use of curricula, availability of professional development, and children's activities while in care. The data will answer such questions as Who are the individuals working in early care and education and what are their experiences in terms of employment characteristics, classroom activities, and professional development?

Questionnaires for each survey are available at:

<http://www.acf.hhs.gov/programs/opre/research/project/national-survey-of-early-care-and-education-nsece-2010-2014>.

#### **D. ADDITIONAL DISTINCTIVE FEATURES**

In addition to the specific questionnaires and sample types embedded within the NSECE design, additional distinctive features of the design are anticipated to be valuable to the policy and research communities.

Many of the specific decisions states face in setting ECE policy depend on issues of price and quality, for example, techniques for strengthening regulatory and quality rating standards; appropriate and effective reimbursement rates and co-pay schedules; and provider organization and individual staff incentives that will produce higher quality in a cost-effective manner. States must also decide how to vary reimbursement levels and program allocations by locality to reflect differences in wages, prices, and available workforce.

ECE providers do not operate in isolation. About 20 percent of children are served by more than one provider, usually one formal and one informal. Children qualifying for part-day, part-year pre-K programs are often enrolled in wraparound or partner services, which are subject to different regulations and have different levels of funding. Effective policies must be informed by knowledge from the family perspective about all services received and costs incurred, and from the provider perspective about the shares of their clients and funding that are related to partnership.

Many communities are trying to develop comprehensive service systems to meet the needs of children and families, relating ECE services to public schools, health care facilities, maternal and child nutrition programs, and family support and other social services.

Minimizing duplication among different financing and operating programs requires that those responsible understand the nature of those services and the staff who work within them. A major impediment to streamlining programs has been uncertainty about whether staff employed in one program can or should be held to the qualifications or certification requirements of another program. The potential to combine the large-scale provider organization survey with a

survey of individual workforce members offers the possibility of providing the comparative information necessary to support strategies for combining or streamlining service programs.

Quality Rating and Improvement Systems (QRIS) are an increasingly common state strategy to improve the quality of ECE. The underlying logic model of QRIS is based on market principles—if parents are given greater information about quality, they will tend to select higher-quality providers and, potentially, pay higher prices for their services; given a combination of supports and financial incentives, providers will make successful efforts to improve their quality. The household and provider questionnaires capture information on parental search behavior, provider admissions and outreach practices, and accreditation and other quality labeling, all of which permit testing of some of the assumptions on which the QRIS model rests.

State and local data systems are being integrated across multiple program domains; many of these include ECE/SA data, and the long-term potential for providing data for research is great. The NSECE can provide starting data and perhaps a model for using administrative data in conjunction with survey data to address policymakers' needs. The NSECE's "51-state" sample not only includes sample in every state and the District of Columbia but also allows research to exploit the tremendous variation in state-level policies such as pre-K availability, licensing requirements, CCDF and QRIS programming, and even variation in macroeconomic conditions at the time of data collection.

Nationally representative research on family child care (FCC)<sup>2</sup> has been seriously hampered by differing state licensing requirements and differing extents of underground provision of services. By capturing core descriptive data on all home-based providers in a consistent manner, regardless of whether they are licensed, registered, or unregulated, the NSECE data provide a full portrait of FCC characteristics and prices and support some investigation as to whether these are consistently related to state regulatory standards, requirements, and enforcement.

The provider cluster design sets up the NSECE data to contribute to the policy-relevant literature on how parents define search areas, providers define catchment areas, and reimbursement rates and other policies may therefore want to define child care markets. A variety of geographic information (such as provider location, household location, and location of parental employment) is incorporated into the NSECE data to inform these geographic questions.

Terminology within the ECE literature has grown organically, often with overlapping meanings or inconsistent interpretations. The NSECE questionnaires have attempted to avoid such terminology wherever possible, adopting lay language and focusing on objective attributes wherever possible. These data have the potential to help clarify appropriate definitions for type of care and other key concepts in ECE research through a flexible approach that will allow equating to a variety of terms used in other data sources.

The NSECE data have been designed to permit linking to complementary ECE data sources, subject to requirements for handling restricted-use, confidential data. Geographically based data sets could be merged: the inclusion of 50 states and the District of Columbia means that researchers can investigate the effects of various policy regimes by linking state- and local-level data such as TANF and CCDF rules databases or state- or local-level licensing requirements.

---

<sup>2</sup> In this document we mostly avoid the common terms family child care (FCC) and family, friend, or neighbor care (FFN), combining all home-based ECE into a pool to be distinguished in analysis. For data analytic reports, we develop and apply new terms on a clearer and more consistent economic basis.

At the provider level, one could attach such provider characteristics as whether or not they accept CCDF subsidies or are NAEYC-accredited. This ability to link greatly extends the value of ACF's investment in the NSECE and leaves open the potential that the data set could be useful in research that could not be anticipated at the time of the study's design or implementation.

As a broad characterization, it can be said that the NSECE design enables capturing the full spectrum and variety of different types of entities, often at the expense of detail about those entities. For example, the instrumentation captures all forms of non-parental care used by households for all of their children under age 13. The sample design emphasizes representing as broadly as is feasible the full range of formal and informal providers of care to children under age 13. Within the provider group, the focus is on characterizing the mix of age groups served or the blend of government program participation. The provider cluster design, together with hooks for integrating local community characteristics, again portrays the heterogeneous interactions of households, providers, local economic conditions, and relevant ECE policies.

## **E. PROJECT TIMELINE**

More than four years of formal design and development work preceded the launch of NSECE data collection. NSECE main study data collection took place from November 2011 through June 2012. To the extent that it was possible, the study attempted to concentrate interviews between January and May 2012; this window, largely free of longer school holidays, was less likely to suffer data quality problems owing to school closures, vacations, providers using "summer" staffing levels, or families reporting summer or alternative child care information.

This summary documents key aspects of NSECE survey design and data collection process. It will be supplemented over time with additional information about data files, constructs, and preliminary analyses to serve as a reference for researchers who will use the project data files or who desire additional background information when reviewing analytical results. Public and restricted-use data files will be made available beginning Fall 2014.