

Texas Child Care Market Rate Survey Study Methodology

Background

45 Code of Federal Regulations 98.45(c) requires that at least every three years as part of the Child Care Development Fund state plan, lead agencies develop and conduct a statistically valid and reliable survey of the market rates for child care services. A market rate survey (MRS) will be considered “statistically valid and reliable” if the study methodology meets the following benchmarks:

Provides complete and current data – The survey uses data sources that fully capture the universe of providers in the priced child care market. In addition, the survey should reflect up-to-date information that is based on price data collected within a specific time period.

Includes the priced child care market – The providers charge parents a price directly that is established through an arm’s length transaction. In an arm’s length transaction, the parent and the provider do not have a prior relationship that is likely to affect the price charged.

Represents geographic variation – The survey includes providers from all geographic parts of the State, Territory, or Tribal service area. It should also collect and analyze data in a manner that links prices to local geographic areas.

Uses rigorous data collection procedures – The survey uses good data collection procedures that includes a response from a high percentage of providers (generally, 65 percent or higher is desirable and below 50 percent is suspect); implements a strong sample design; conducts analyses of potential response bias to ensure that the full universe of providers in the child care market is adequately represented in the data and findings; and survey based on the languages used by child care providers.

Analyzes data in a manner that captures market differences – The survey examines the price per child care slot so that all child care facilities are not equally weighted because some serve more children than others. The survey also collects and analyzes price data separately for each age group and category of care to reflect market differences.

Texas Workforce Commission Child Care Market Rate Survey Methodology

TWC contracts with the University of Texas at Austin School of Social Work and the Ray Marshall Center to conduct an annual Texas MRS using a methodology that complies with the required federal benchmarks. The TWC methodology creates statistically valid and reliable rate distributions of part-day and full-day child care market rates by:

- Geographic area (i.e. for the entire state of Texas and for each of the 28 Local Workforce Development Areas (LWDAs);
- Provider facility types (i.e. licensed child care centers (LCCCs), licensed child care homes (LCCHs), and registered child care homes (RCCHs); and

- Age groups (i.e. infants (0-11 months and 12-17 months); toddlers (18-23 months and 2 years); preschoolers (3 years, 4 years and 5 years); and school-age children (6-13 years)).

The survey also has rigorous data collection and rate estimation processes, adequate response rates; is available in Spanish and Vietnamese; and weights rates by the number of slots the rate represents.

Data Sources and Sample Population

The MRS sample comes from a list of all licensed and registered child care facilities provided by the Texas Health and Human Services Commission Child Care Regulation (HHSC CCR) and the TWC administrative database of providers with agreements.

The MRS study sample is drawn from the list of all licensed and registered child care facilities provided by HHSC CCR generally about a month before the surveys begin. The surveys are typically conducted between September and March of the study cycle which captures complete and current information for the rates estimates.

The list of facilities from which the sample is drawn is typically between 12,000 and 15,000 and includes the three types of child care providers (LCCC, LCCH, RCCH) in each of the 28 LWDAs.

TWC Administrative Data

As has been done since the 2018 report, the child care daily market rates are estimated by including both data from the survey and data from administrative records maintained by TWC, known as published rates. These published rates are subjected to several screens, then special weights are calculated so that their share of care provided is not be overstated.

A database of published rates is provided by TWC in July, from which published rates of subsidized facilities are drawn. Rates from facility types other than LCCCs, LCCHs, or RCCHs are excluded from further analysis. Only rates that were indicated as active rates as of August 15th are included, which eliminates redundant rates from the same provider. Also excluded were rates for anything other than “regular day care,” which eliminates night and weekend rates. Finally, any part-day rates that were as high or higher than the corresponding full-day rates from the same provider and age group were removed, as these are not true part-day rates.

Before the published rates can be combined with those from survey data, it is necessary to remove duplicate rates from facilities that are surveyed. Finally, a small number of rates from facilities who could not be linked to the survey or whose capacity was unknown are removed. For remaining rates, the number of slots attributed to each rate is estimated by a statistical model utilizing the facility’s reported capacity, the typical share of slots allocated to each age group and full or part day status, and the estimated share of slots dedicated to Child Care Services (CCS) children in facilities of this type.

Determining the Priced Child Care Market

The population of interest in this study is the priced child care market in Texas. Licensed and registered facilities that charge families directly and do not discount rates based on prior relationships with families are included in the survey to determine child care market rates. Published rates of CCS providers are also included by incorporating administrative records data, with weighting schemes to ensure that CCS facilities carry as much weight in the final analysis as the share of child care they represent

The following providers are excluded from the list of available providers prior to the sample being drawn or at the time of the survey interview:

- Head Start facilities (do not charge families directly);
- Drop-in care facilities (charge very high hourly rates and not typically used by employed parents for regular care);
- part-day care with no after-school care;
- summer camps;
- care provided to specific populations only (i.e. children with special needs, children of teen moms, children of staff at a company, etc.);
- free/family-discounted child care services; and
- school and kindergarten programs that did not offer regular afterschool care.

Sample Size and Selection Methodology

After the facilities identified prior to the interview are removed from the statewide universe of facilities, a small statewide representative sample of 1,000 facilities is randomly selected in order to:

- 1) provide greater statistical power for estimating statewide rates and trends; and
- 2) increase the precision of rate estimates for the largest local boards that provide the bulk of care in the state.

In addition to the statewide sample, a second random sample is drawn from the remaining facilities to estimate rates at the LWDA level. For the second sample, a target number of centers and a separate target for homes from every LWDA based on projections of the number needed to yield an average of 200 independent rate observations across full-day rate categories.

Projections based on detailed response rates to recent annual surveys indicate that, on average, 66 LCCCs, 33 LCCHs, and 79 RCCHs per LWDA are needed to reach this goal. These target numbers are then increased by between 3 percent and 7 percent for LWDA and facility type combinations that showed low response rates in recent surveys. The target number of facilities is then randomly selected from each LWDA, or if fewer providers are available in any given combination of LWDA and facility type, the entire population is selected (a complete census).

Identify TWC Providers in the Sample

Before the survey sample is finalized, a match is performed between the list of sampled facilities and a database indicating which facilities are serving CCS children, whose published rates are already available through administrative records. These partially subsidized facilities are randomly split into two groups, with one half being retained in the sample and the other half set aside to be excluded from further study. This sampling plan can be described as an under-sample

of subsidized facilities, whose rate data are largely redundant with existing data sources, and a corresponding over-sample of non-subsidized facilities, whose rates constitute the bulk of the paid market and are not otherwise available without surveying.

Before final rates are estimated, data from the statewide sample are combined with that from the local samples in their respective areas, as well as the published rate data, and weighting schemes are applied as appropriate. The number of rate observations collected by this design is more than adequate to provide stable rate estimates for all LWDA, without needing to rely on other estimates or approximations.

Survey Data Collection Standards and Process

Data Collection Process

After the sample is determined, surveys are conducted by telephone interview between September and May.

Prior to interviews, the research team mails introductory postcards and/or emails to all facilities in the sample explaining the survey's purpose, objectives, confidentiality policies, and the voluntary nature of participation. Information is presented in both English and Spanish unless providers' surnames suggested they are of Vietnamese origin, in which case information is presented in both English and Vietnamese. A website is also available for providers to obtain further information about participating in the survey.

The Market Rate Surveys for home-based and center-based facilities cover similar topics. The major difference is that home-based facilities provide detailed information concerning each child under their care (age, exact schedule of care, and rate charged), whereas center-based facilities provide rates by age group and by whether care is full-day or part-day (less than six hours per day).

The survey collects information on:

- hours of operation;
- vacation schedules;
- child age and enrollment;
- rates;
- additional fees (i.e., registration, enrollment, activity);
- administrative categories (i.e., non-profit/for-profit, associations, transportation)
- national accreditations;
- Texas Rising Star program; and
- serving children and youth with disabilities or special health care needs.

Data Collection Standards

Detailed training and supervision are provided to research staff members on survey methodology, interviewing protocols, data collection, data entry and tracking procedures by the project manager and more tenured research staff.

Because different types of facilities tend to set their rates differently, analysis of center rates is based on the rates for a category of care (for example, full-day care for toddlers), while the rates for home-based care (LCCHs and RCCHs) are determined by the amounts charged for individual children. In interviews with centers, the researchers ask for the rates charged for a particular age group following a particular schedule. In contrast, researchers ask home-based facilities for the age and rate charged for each individual child.

After the survey data are collected, all rates that appear extremely high or low (known as outliers) are identified, and researchers individually check each rate report and correct any errors they found. Subsequent to these corrections, a small portion of the most extreme outliers remaining in the sample are corrected to remove their excessive influence on statistical measures of the rates distributions. Similar procedures are implemented to detect instances in which the number of slots reported for a given rate is too extreme, whether too high or too low, and to correct these extremes to more reasonable values. In general, as one might expect in a large and diverse state, extreme rates, whether at the low or high end of the scale, are typically the result of unusual services provided by an individual facility and/or variations in subsidies and financing rather than research error.

Response Rates

The overall response rate is calculated based on the eligible child care providers that completed the survey.

Response Bias

In order to adjust for any bias due to what types of facilities responded to the survey, a response model was developed using logistic regression. Predictors for this regression primarily include measures derived from registry data maintained by HHSC CCR, as well as county-level measures from another study. The HHSC CCR measures include indicators for self-reported features of care provided including offers transportation, accepts subsidies, serves infants, toddlers, preschoolers, school age, special skills, field trips, get well care, as well as after school, drop in, part time, or weekend care. Additional measures include the facility age in years, based on the initial license issue date, as well as age-squared to account for potential curvilinear effects. The dependent or response variable for this regression is an indicator showing whether the facility completed the survey, in whole or in part. Facilities determined to be ineligible for the survey are excluded from the regression.

If a bias exists, weights are computed by inverting the estimated probability of response for individual facilities and are included in analyses to allow adjustment for this bias at either the facility or rate level.

Analytic Methodology and Levels of Analysis

Project researchers summarize the survey data and conduct analysis at several distinct levels. The most basic analyses are done at the level of individual providers. These results document the proportion of providers responding to the survey that offer various services (e.g., non-traditional hours, transportation), or that fall into different administrative categories (sectarian or for-profit). The exception to this rule is the one item that measures the “percent of enrollment that is subsidized,” which is analyzed at the level of the individual child. A similar item concerns the percent of facilities that serve any subsidized children. When this number is higher than the “percent of enrollment” number, it indicates that the subsidized children in an area are widely

dispersed among facilities, and when it is lower, it indicates that the subsidized children are relatively concentrated among a smaller number of facilities in the area.

The remaining analyses are conducted at either the rate observation level or at the level of the child care slot. Each center or home-based facility can contribute more than one rate observation to the analysis, and each rate observation can represent more than one slot. However, the nature of this relationship depends on the type of facility, as described below. The estimation of market rates includes a combination of rates from this survey and administrative records regarding published rates of subsidized providers, so different procedures were also followed for administrative records.

Among survey data gathered, daily market rates for licensed centers are captured for all categories of care offered, regardless of whether any children were currently being served in such categories. The categories consist of all possible combinations of age groups (gathered for actual age categories in which each center offered rates but aggregated to the eight age categories for reporting) by Full-day status (part-day or full-day). Thus, one center could contribute as many as sixteen independent rate observations, each representing any number of children (including zero). These rate observations are then weighted by the number of child care slots they represent when calculating market rate distributions.

In contrast, market rates for surveyed home-based facilities, including registered homes and licensed homes, are gathered at the level of the individual child currently being served. Each child's age, detailed weekly schedule, and rates charged were gathered for purposes of calculating daily market rates. In this case, the individual children are treated as independent rate observations, and each facility could contribute as many as nine observations (or fewer depending on the facility type and number of children enrolled). Because of this, there is no need to differentially weight the rate observations by the slots they represent when calculating the local market rate percentiles for surveyed home-based facilities: each child (or rate) starts with a weight equal to one.

For the analysis of surveyed center rates, the number of child care slots for each rate category is determined in one of two ways. First, for categories of care in which children are currently being served, the number of children in each category served as a proxy for the number of slots. Second, rate categories in which no children are currently being served are also included in the analysis, since they were also theoretically part of the market. This is done by estimating the number of slots for each of these rate categories with a formula that multiplies the number of children served at each facility by the average proportion of children, across all licensed centers, served in each rate category. Thus, for example, a facility that served 100 children and had an existing part-day infant rate schedule but did not currently serve any part-day infants, would have its number of slots for this category of care estimated at two ($100 \text{ child capacity} \times 2 \text{ percent of children served in the part-day infant category across all licensed centers}$). If the same facility served no part-day preschoolers, its number of slots would be estimated at seven ($100 \text{ child capacity} \times 7 \text{ percent served in this category across all licensed centers}$). This method allows fuller use of the rate information gathered from surveyed centers, especially for rare forms of care in which rate observations are otherwise scarce.

Statewide Estimation

For statistics conducted at various levels of analysis, special procedures are necessary to compute statewide versions of the tables. Since less populous areas may be over-sampled for the survey (meaning that facilities in these areas are more likely to be selected for the survey), a correction is applied to estimate statewide rates without over-representing facilities in these smaller areas. Starting with the 2018 survey and continuing unchanged to present, subsidized facilities are under-sampled in order to concentrate survey resources on non-subsidized facilities whose data are more difficult to gather, while amending data on subsidized facilities by adding administrative records data. The correction for both these factors involves computing weights that, when applied to the statewide analysis, equalize the probability that any given facility or rate might have been selected at random.

The calculation of weights for statewide estimation varies slightly depending on the level of analysis: facility level or rate level. For facility-level analyses (as in the Services and Enrollment and Administrative Categories sections), the weight is the inverse of the probability of selecting the facility at random. Thus, for less populous areas in which all facilities are selected, the weight would equal one. For more heavily populated areas in which, for example, 20 percent of facilities are selected, the weight would equal five (1 divided by 0.20). When these weights are applied to any given statistical analysis, the facility with a weight of five would count five times as much as the facility with a weight of one, and the result would be representative of the state as a whole. Calculation of statewide rates is similar to estimation of local rates in that each rate is weighted according to its share child care slots.

Small-Area Estimation

In many less populous LWDA's and for some exceptionally rare rate categories (particularly part-day infant and toddler care in homes) the number of rate observations would have been too small to estimate stable rates using traditional methods, even for the many areas in which the entire universe of facilities had been sampled. For this reason, an estimation methodology first developed for the 2010 survey is used. Using this approach, parameters are estimated that fully characterize the distribution of rates in each category of care. These parameters are then applied, using an assumption of normality of the rate distributions, to estimate each rate percentile of interest. The primary advantage of this solution is that a medium to large sample of rates is utilized to estimate each parameter, thus the estimates based on this methodology are more stable over time, and they fall into the expected patterns more frequently.

Prior to any estimation of parameters, the daily rates themselves are first transformed to make them assume a more normal distribution. This transformation is done by taking the positive square root of the daily rate. The effect of this is to minimize the influence of high-end outliers on estimates of the parameters of distributions. This transformation is reversed later, following all estimation, by squaring the rates. The next step in the estimation of parameters involves estimating area effects, the average rate for each LWDA, while statistically controlling for all other features of the rates involved in the estimation (including facility type, age group, and rate type).