COMMON CHALLENGES IN THE STUDY OF CONTINUITY OF CHILD CARE SUBSIDY PARTICIPATION

Methodological Brief OPRE 2012-55
December 2012

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.

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.

ACKNOWLEDGMENTS
Funding for this study was provided by grant 90YE0119 from the Office of Planning, Research and Evaluation, Administration for Children and Families. Our work has benefited greatly from the research conducted with the other members of the Five-state Child Care Subsidy Dynamics team, particularly Deanna Schexnayder and Daniel Schroeder, and from feedback from Child Trends and members of the Child Care Policy Research Consortium.

This Brief was developed by members of the Child Care Policy and Research Consortium's Child Care Subsidy Workgroup, which is supported by the Office of Planning, Research and Evaluation through the Child Care and Early Education Policy and Research Analysis and Technical Expertise contract with Child Trends.

Common Challenges in the Study of Continuity of Child Care Subsidy Participation
CCPRC Subsidy Workgroup Methodology Research Brief Series

OPRE 2012-55
December 2012

Submitted to:
Ivelisse Martinez-Beck, PhD., Project Officer
Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Submitted by:
Elizabeth E. Davis, University of Minnesota
Deana Grobe, Oregon State University
Roberta B. Weber, Oregon State University

Contract Number: GS10F0030R
Activity Director: Nicole Forry
Child Trends
4301 Connecticut Ave NW
Washington DC, 20008

 Suggested Citation: Davis, E. E., Grobe, D., & Weber, R. B. (2012). Common challenges in the study of continuity of child care subsidy participation. Methodological Brief OPRE 2012-55. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

UNIVERSITY OF MINNESOTA
Driven to Discover℠

Abstract

In this paper we discuss several key challenges encountered when conducting a study of the continuity of participation in the child care subsidy program. While many of these issues are familiar to those who have studied participation dynamics in other assistance programs, and to those familiar with survival analysis, we describe these challenges and solutions in the context of the child care subsidy program. This brief is intended to help grant reviewers, policymakers, and researchers new to the study of dynamics of subsidy participation when they plan, conduct and evaluate studies of child care subsidy participation. Ignoring these issues can lead to study results that are not comparable with other studies and may yield misleading findings. Given the importance of child care subsidies for low-income families and children, there is a need for comparable research that increases our understanding of families' experiences with child care subsidies and the influence of policy differences across states on the continuity of subsidy participation.

1. Introduction

Child care subsidies funded through the Child Care and Development Fund (CCDF) are intended to support both parental employment and children's development in preparation for school. Subsidy participation is more likely to support both of these goals when subsidized child care arrangements are stable and high quality. However, most of the states that have measured spells of subsidy participation find relatively short periods of participation and frequent returns (Meyers, et al. 2002). Short spells of subsidy use may result in frequent changes in child care arrangements (Weber, 2005; Ha, 2009). Data on how long families have been receiving subsidies provide important information about the families' experience of the program and thus whether or not the program is likely to meet its goals. These findings also have fiscal implications for the cost of the program. Yet answering questions about the dynamics of program participation is not as straightforward as other types of policy evaluation or program impact studies. The purpose of this brief is to describe the challenges in studying the continuity of subsidy participation and some of the preferred methods for answering questions about subsidy participation dynamics. The goal is to provide guidance to help grant reviewers, policymakers, and researchers new to the study of continuity of subsidy participation as they plan, conduct and evaluate studies of child care subsidy participation.

Studying the length of time until an event occurs typically involves statistical methods called survival analysis or event history analysis. The basic outcome of interest is the length of time of child care subsidy participation (duration). Although consensus on how to measure duration has emerged in the fields of welfare and child welfare research, no such consensus exists in the field of child care and early education. Prior research on continuity in the child care subsidy program has used inconsistent methods. This makes comparisons of findings across studies extremely difficult, especially in cases in which the researchers do not describe their methods in sufficient detail. Studying continuity in the subsidy program is feasible and being able to compare findings across states is desirable. Reaching consensus on duration study methods will support cross-state studies and thus will substantially advance knowledge of continuity in the subsidy program.

Information about patterns of child care subsidy use can assist both federal and state policymakers in designing and developing child care subsidy programs. For example, better understanding of the impact of policy interactions on the people served will allow states to decide which families they want to target for the child care subsidy program. Information on participation dynamics may also guide development of policies to impact the continuity and stability of child care arrangements.

One approach to learning about families' experience with the child care subsidy program and the continuity of their participation is to survey parents. Surveys can (and have) asked parents about current and past subsidy participation. Parent report of current subsidy receipt may be reasonably accurate (Herbst and Johnson, 2010), but asking parents to recall past subsidy receipt is likely to be more error-prone. In addition, collecting enough data to track parents' subsidy participation over time would be difficult in a survey. Thus, most studies of the dynamic patterns of subsidy participation use longitudinal data on subsidy receipt obtained from state administrative data systems. Administrative data systems provide a consistent measure and a large number of time units (weeks or months) in which to observe families. One disadvantage of administrative data is that they tell us little about the reasons why families leave the subsidy program, or why they return. Interview and survey data can fill in the gaps left from administrative data. They provide answers to questions such as why a parent exits or what happens to a child care arrangement when a subsidy spell ends.

A caution in using administrative data is that they are collected by systems designed to determine eligibility and process payments; these systems are not designed for research purposes. Because of this, an administrative dataset commonly lacks documentation on the quality of the data and often lacks important information that can be used as control variables (for example, education and prior employment history). Furthermore, the data are typically available only for periods that the family participates in the program, limiting what we know about what happens when they leave (which might impact the likelihood of returning to the program). Despite these disadvantages, administrative data are a convenient and inexpensive source of child care subsidy participation data.

In this brief, we assume that the data available are from a state administrative database on subsidy participation and are based on calendar months. Most states report the number of children and families receiving subsidies to the federal government on a monthly basis. However, state subsidy data systems may use weekly or biweekly periods of time for authorizations and payments. One could use a unit of time other than a month in the type of analysis described in this brief. However, we assume monthly data are the most widely available, and use of a common unit of time will support comparability of findings across states.

DEFINITIONS

Throughout this brief, the following distinctions are made for the terms "spell," "spell-length" (duration), and "dynamics":

A **spell** of subsidy receipt is defined as the number of consecutive months a family (or a child) received a subsidy without a "break"; that is, a month (or more) in which the family (or child) did not receive a subsidy.

Spell length measures the length or duration of a spell of participation, which is the number of months a family receives a subsidy for child care without interruption (continuous subsidy receipt). The duration or length of the spell is usually measured in months but could be measured in other time units.

Dynamics refers to various time-related elements of participation in the child care subsidy program (for example, spell-length, total number of months of participation, time between spells, etc.). **Continuity** is also used to refer to the duration of subsidy participation as well as the frequency of exits followed by quick returns to the program.

2. How to begin: Data structure and defining a spell

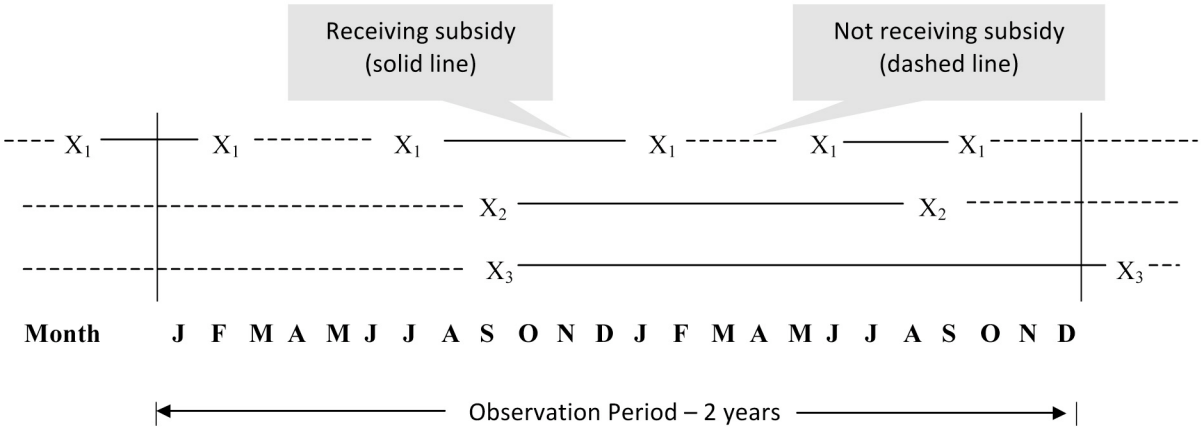
This brief focuses on the technical issues around studying dynamics of program participation. We start with a discussion of the data used in the analysis of participation in the subsidy program over time. Understanding the data needed for this type of study and the challenges in creating the proper data structure and sample are key to producing reliable findings.

2.1 What do the data look like?

For an analysis using event history or survival analysis methods, we need longitudinal data (that is, data over time) on the occurrence of events. An event, in this case, is a change in program participation, for example, a child receiving subsidy no longer does so. We are interested in *when* the event occurred, that is, when the child started and stopped receiving a child care subsidy. The data structure most frequently used in this type of study is in the form of monthly observations with an indicator for whether or not the child participated in that month. Sometimes the data for a child exist only if the child participated in that month (and is missing or blank otherwise). Alternatively, the data can be in the form of a start and end date for subsidy program participation.

The figure below illustrates monthly data on program participation and the definition of a spell. If we consider a two-year observation period, the pattern of subsidy use for three children (X_1, X_2 and X_3) might look like Figure 1.

Figure 1: Pattern of Subsidy Participation (Monthly Observations)



In Figure 1, two X's with a solid line between them denotes the occurrence of a "spell" of participation, dotted lines represent a period of non-receipt, and the solid vertical lines indicate the beginning and end of the observation period. Thus a spell of subsidy receipt is a continuous stretch of consecutive months in which a family received a subsidy without a "break" (a month or more) in which they did not receive a subsidy. In the figure, the first spell for child X_1 began prior to the start of the study's observation period as shown by the leftmost X_1 that is to the left of the vertical line. The period of participation for child X_3 ends after the end of the observation period. Later in the brief we will discuss how to deal with the problems caused by start and end dates outside the observation window.

The selection of a length of the observation period is an important study design decision. Ideally one would want to have a long enough observation period to capture the range of experience in the program. If most children receive subsidies for more than one year, then a one-year observation period would be too short to reflect their experiences accurately. On the other hand, using a very long observation window such as more than five years may have drawbacks as well. Policy and economic changes during the study period may confound results in unknown ways.

Using administrative data about children and families in the child care subsidy system usually requires merging data files that are at the child-level with data at the family level (and sometimes also at the provider or arrangement level). It is very important to check that these matches are done accurately. Grobe, Weber and Davis (2003) provide some advice on merging child, family and provider data files.

2.2 Defining a spell

The basic idea of a spell is intuitive – a period of time during which the child continuously received subsidized child care services without interruption – but the detail of measuring subsidy spells is a bit more complex. We can see some of these issues in Figure 1, for example, start and end dates outside the observation window and multiple spells within the observation period. Other critically important issues relate to how to measure the receipt of subsidized services – at the child or family level -- and how to define the end of a spell when the child returns after a brief hiatus. These issues are discussed next.

Defining subsidy receipt and the end of a subsidy spell[1]

With administrative data from a state, defining receipt of subsidized child care services may seem straightforward. However, in order to accurately measure the length of time a child participates in the subsidy program, participation must be based on when the care was received, not when payment was made. Federal reporting requires that states base their reports of subsidy participants on service dates, not payment dates. For example, if a child received care from a child care provider in May, June and July, and the provider was paid for this care by the subsidy program in July and August, the child's subsidy participation spell would include the months of May, June and July.

The end of a spell of subsidy participation is indicated by a period of time without subsidized services. The question is, however, how long a period of time is needed to indicate an interruption in subsidized care. A few days absent for illness or holiday would generally not be considered a break in subsidized care. In the literature on welfare participation, researchers typically use a three-month break to define the end of a spell of welfare use in order to not interpret churning due to paperwork problems as an actual break. However, unlike the subsidy program, parents who receive their welfare check late still get the assistance. If the family needed the care services and the child did not receive them for an entire month, then there was a real interruption in subsidized care. The key is that the data on subsidy participation must be correctly defined based on service dates – when the care was provided – regardless of when payment occurred. When using data based on service dates, most researchers have used a one-month break to define the end of a subsidy spell.

[1] Note that we are interested in the receipt of subsidized child care services, not the period of time that a family is eligible to receive services.

It is worth testing whether a different definition of a break to define spells alters one's results. Table 1 below shows the estimated spell lengths for one state using two different definitions of the end of a subsidy spell. In this particular state, redefining a break in subsidized care to last for two months instead of one made a small difference in the estimated median spell length (and other percentiles of the distribution). In contrast, Ha's (2009) study with Wisconsin administrative data used both one- and two-month breaks. With a two-month break the median duration was 8 months compared to 6 months with a one-month break between spells.

For any particular state, comparing the results with a one- versus two-month break may be important. However, it is important to keep in mind that if the state data system properly records service dates and spells are based on when subsidized care was received, rather than when it was paid for, then any gaps in receipt are true periods with no subsidized child care. If families are returning within one or two months at a high rate, policymakers may need to review policy to determine why. Witte and Queralt (2005) found that about 20% of those who returned to subsidy did so after only one month, and using a two-month gap to define a subsidy spell end resulted in an increase in the median spell length of one month.

Table 1. Comparison of One and Two Month Breaks in Subsidy Use (An Example)

Spells defined using one-month break		Spells defined using two-month break	
	Spell length (months)		Spell length (months)
25th percentile	3.3	25th percentile	3.7
Median	7.4	Median	8.5
75th percentile	16.9	75th percentile	19.4

Data source: Actual child care subsidy data from a state administrative data system for 2005-2010.

Technical note: Spell duration percentiles were estimated using an accelerated failure time (AFT) model assuming a log-normal distribution and adjusting for right-censored spells that have not been completed by the end of the study period.

Child versus family level spells

Depending on the research question of interest, subsidy participation may be measured at the family-level or the child-level. A child receives the subsidized care and a child spell is defined as continuous receipt for an individual child. If the family has more than one child receiving subsidized care, one child may stop while another continues to receive subsidized care. From the perspective of certain family decisions, it may be appropriate to consider using family-level spells. A family-level spell is defined as a continuous receipt of a subsidy for any child in the family; it begins when at least one child is participating and ends only when no child is participating. The appropriate unit of analysis will be related to outcomes of interest. For instance, if the outcome is parental employment, using the family as the unit of analysis makes sense whereas if the study focuses on child outcomes (e.g., child care arrangement stability), a child-level analysis of spells would be the logical approach.[2]

[2] The relationship between family spells and child spells (i.e., the subsidy spells of children in the same family) is an under-researched topic.

If the decision is made to base the analysis on child-level spells, another methodological issue is whether to include all the children in a family as separate units of analysis. Two concerns arise: first, that the analysis will overweight the experience and characteristics of families with more children, and second, that the decisions (such as exiting subsidy) will be highly correlated for children from the same family (thus they are not independent observations). While these issues can be dealt with in the analysis, an alternative approach is to pick one child to represent the family. This child may be selected randomly (preferred approach) or the youngest or oldest child could be chosen to represent the family.

3. Who to include in the study and censored spells

3.1 Censoring (missing information on the start or end date of the spell)

When we observe spells of subsidy participation during a specific period of time, a family's spell may begin before or end after the observation period. Left censoring describes cases in which a spell begins prior to the study date, and right censoring refers to cases that continue past the end of the observation period. Thus there is often missing information on the start date of on-going or left-censored cases, and also missing information on the ending date of the right-censored cases (because they do not end within our window of observation).

Referring back to Figure 1, we see that the first spell for child X_1 began prior to the start of the study's observation period as shown by the leftmost X_1 (which is left of the vertical line) and thus this spell is left censored. The spell for child X_3 ends after the end of the observation period and thus is right censored. The two spells occurring during the observation period for child X_1 and the spell for child X_2 are uncensored.

Left-censored cases present two problems for the analysis. First, one may not know the start date of the spell of participation because the case was ongoing at the beginning of the study. Treating the left-censored spell as if it began at the beginning of the observation period will introduce bias because the researcher does not know when that child began receiving subsidy for that spell. However, even if the start date is known, including the left-censored cases introduces "length bias" because children who started at the same time but exited prior to the observation window will not be included in the sample (Swartz, et al. 1993). Including only those children whose subsidy continued into the observation window will bias the findings to longer spell lengths.

The most common approach is to discard all left-censored cases, that is, to exclude all the months in any spell that is on-going in the first month of the observation window. If the child has another spell of subsidy receipt, one that begins during the observation window, that spell should be included in the study.

Although removing left-censored cases is a common solution, one should understand that this method may exclude those with the longest spells, so the longest spell-lengths could be underestimated.[3] For reasons discussed further below, dropping the left-censored cases is the preferred strategy as long as the length of the observation window is sufficiently long.

"Right-censored" spells are those that have not been completed by the end of the study period. In Figure 1 above, the spell for X_3 began during the observation period, but ended after the observation period ended. Fortunately, these cases can be included in the study if one uses particular statistical methods (described below) that account for the right censoring.

[3] Certain statistical methods allow the inclusion of left-censored observations in the estimation, and methods can be used to test for statistical association between spell-length and other variables (see Allison, 2011). Swartz, et al. (1993) tested whether excluding left-censored cases led to underestimation of spell lengths (for those lacking health insurance coverage) and found no difference in results.

3.2 Defining the study population: Entry cohort versus point-in-time approach

Analysis of participation data over time is different from studies that use cross-sectional data (at a point in time). The unit of observation for a duration study is a month (or other time unit) of participation rather than a family or child. Some families and children have more observations (months) than others. Special attention must be paid to issues related to who is included (and excluded) in the study even when using the full population of subsidy participants from state administrative data records.

The key sample design issue in studies of dynamics is how to capture the experience of all participants regardless of whether they participate for a few months or several years. This issue arises whether the study includes all program participants or is based on a (random) sample. Many program managers and policymakers typically think about the families who are participating in the program at a point in time, for example, those who are on the rolls this month. Thus, they tend to pose questions about program dynamics in terms of the current caseload: how long have these families been receiving subsidies, and how much longer will they stay on? However, the families participating at a point in time, e.g., the current month's caseload, are not representative of *all* families who participate in the program.

This issue arises because some families[4] have more observations (months of participation) than others. An analysis based on the current caseload (a point-in-time sample), will oversample those families with more months of participation. To analyze the experience of all families who participate, an *entry cohort* approach is more appropriate.[5] The entry cohort approach includes all families beginning a spell during a specified time period, typically a year or number of years. As discussed earlier in the section on censoring, all months of a left-censored spell for a family that is on-going at the start of the observation period are dropped from the analysis. Families may still be receiving a subsidy when the study ends, however, the methods to estimate spell length can take into account this right censoring problem.

It makes a huge difference whether one uses the entry cohort or point-in-time approach. Two examples below illustrate this point. The first set of figures show the different distributions of spell lengths in two different states (labeled A and B) under the two approaches. Figures 2a and 2b compare the estimated spell length distribution for all spells that were in progress at a point in time in the middle of the study period and for all spells that began during the five-year period (entry cohort). Of all spells that begin in State A during the study period (the entry cohort approach), more than half are quite short, lasting three months or less. At a point in time, however, the families currently receiving a child care subsidy are likely to be in the midst of a spell that will last ten months or more on average. In State B, the results are similar, though spell lengths on average are longer. Fewer than 40 percent of new spells are three months or less in State B, yet nearly 20 percent of children on the caseload at a point in time are in the midst of a spell of three years or more.

[4] This paragraph assumes that the analyst is using family-level spells but is applicable for child subsidy spells as well.
[5] Statisticians sometimes refer to point-in-time as stock sampling, and entry cohorts as flow sampling. For an example using welfare caseload data, see Klerman and Haider (2001).

Figure 2a. Distribution of subsidy spell length for both entry cohort and point-in-time spells for State A.

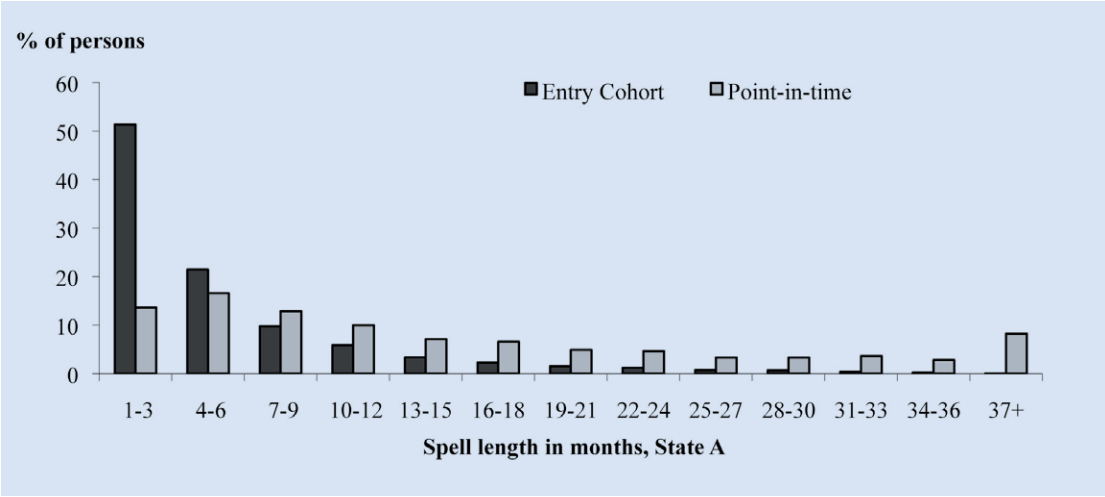
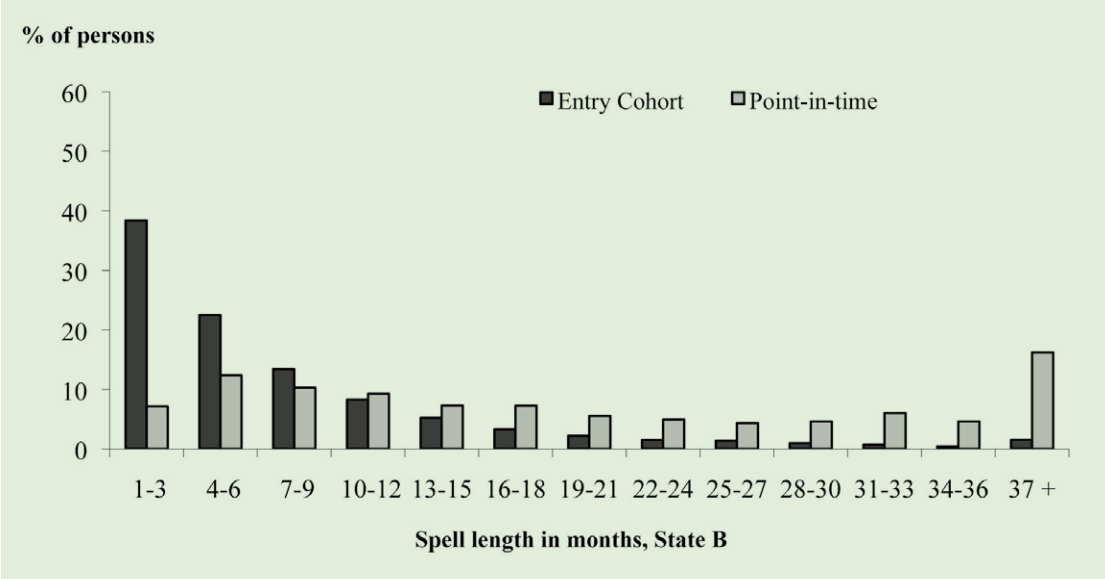


Figure 2b. Distribution of subsidy spell length for both entry cohort and point-in-time spells for State B.



Key message

It may seem counterintuitive, but the following two statements are both true. In State A (using actual state administrative data):

i. The median length of stay for a family was three months. In other words, half of all families who start receiving a subsidy left the subsidy program in three months or less.
ii. Of the families receiving subsidy in a particular month, half received a subsidy for 10 months or more.

The first statement is based on the entry cohort approach while the second relies on a point-in-time sample. Using the entry cohort versus the point-in-time approach will yield different conclusions. The entry cohort approach is most commonly used and is more appropriate for most studies of duration of subsidy participation.

Table 2 offers a second illustration of the difference in findings that results from using a point-in-time versus entry cohort approach. In Table 2, the percentiles of the spell distributions were estimated first for all families with a subsidy spell underway at a point in time in the middle of the observation period, and second for the first observed spell for all families beginning a spell during the five-year observation period (entry cohort approach). In both states, the estimated median spell length was much longer based on the point-in-time approach compared to the entry cohort. For State A, the median spell length was 3.9 months using the entry cohort compared to 11.3 months using the point-in-time families. For State B, the difference was even larger, 5.7 months versus 16.5 months. This discrepancy occurs because the point-in-time approach oversamples families with long spells of participation. Similar differences occur at other points of the spell distribution such as the 25th and 75th percentiles.[6]

Table 2: Difference in Estimated Spell Length Distribution, Point-in-time versus Entry Cohort Approaches

	State A		State B	
Spell length duration (months)	Point-In-Time	Entry Cohort	Point-In-Time	Entry Cohort
25th percentile	6.0	2.0	8.1	2.8
50th percentile (median)	11.3	3.9	16.5	5.7
75th percentile	21.5	7.6	33.4	11.4

Data source: Actual child care subsidy data from two different states' administrative data systems.

Technical note: Spell duration percentiles were estimated using an accelerated failure time (AFT) model assuming a log-normal distribution and adjusting for those "censored" spells that have not been completed by the end of the study period. Point-in-time sample included all families receiving subsidy in one selected month. Entry cohort sample included the first observed spell of each family who began a spell of subsidy receipt during a five-year period.

For many purposes, the entry cohort approach is used in order to analyze the experience of all families in the program and to avoid over-representing the families with longer spells.[7] For certain policy questions, however, analysis of the caseload at a point in time is more relevant. For example, the ongoing cases at a point in time receive a greater share of program resources than do all families, because of their longer spells. Thus, researchers may choose to over represent these longer-spell families in their analyses when their research questions are concerned with resource allocation. It is critically important, however, that one recognize that estimated duration of spells will be longer using this approach.

[6] The difference across the two states is clear with either approach: subsidy spells are shorter on average in State A than in State B. One possible reason for differences in median spell length across states is differences in which families are participating in the program. An earlier study found that subsidy spells are shorter for families receiving subsidies while in job readiness or training programs compared to families who are employed (Meyers, et al., 2002).

[7] Researchers in other fields have compared results including both the left-censored (on-going spells) and entry cohort spells. To do so requires more complex statistical methods – simply estimating spell lengths including both is not appropriate. If the data include the start dates for the left-censored cases, one can adjust the likelihood function to account for the bias of including these longer spells. See Schwartz et al. (1993) and Berger & Black (1998) for examples. Most social science researchers simply drop the on-going (left censored) spells from their analysis.

3.3 Multiple or repeat spells

Referring once again to Figure 1, we see an example of a child (X_1) who has more than one spell of subsidy receipt within the two-year observation window. There are three spells with breaks of at least one month between them for child X_1, including a left-censored spell. Most studies of participation spells include only one spell per child (or family, depending on the unit of analysis), typically the first one that begins during the observation window (and therefore is not left censored). Analysis can be done on multiple spells, however. Without going into the technical details here, the length of the second spell is likely to depend on the length of the first spell. Those with multiple spells tend to have shorter spells. One alternative is to use a randomly selected spell for those with multiple observed (non-left censored) spells within the observation period rather than the first spell. The most common approach, however, is to use only the first spell that begins during the observation period for each family (or child).[8]

3.4 Completed Spells

Note that some researchers consider using only completed spells, that is, spells for which both the start and end dates are known. This approach seems appealing because it appears to eliminate the problems of both right and left censoring.[9] However, this method is likely to lead to biased results. As noted above, including the left-censored spells oversamples the long spells and so typically these are excluded from the analysis of spell length. But excluding the right-censored cases would create a bias in the sample toward shorter spells because those who continue beyond the observation period are likely to have longer spells and their spells would be excluded. Since statistical methods can account for right-censored spells, there is no need to exclude them from the analysis.

In Table 3, we have taken the information from Figure 1 on the three children's subsidy spells and presented it in table format, showing the start and end date of each spell for each child. In this example there are five observed spells of subsidy participation. Three are completed spells. Note that the length of any completed spell can be no longer than the observation window. Completed spells are not a random subsample of all spells, and those who have not left the subsidy program during the observation period are the least likely to do so. Thus a study that uses only completed spells of subsidy participation in the analysis will produce biased estimates of the duration of spells.

Table 3: Example of Subsidy Participation Spells Data Assuming a Study Period of January 2008 to December 2009

	Spell number	Month that subsidy began	Month that subsidy ended	Length of spell	Characteristic of spell
Child 1	First spell	Oct. 2007	Feb. 2008	5	Left-censored
Child 1	Second spell	Aug. 2008	Jan. 2009	6	Completed
Child 1	Third spell	Jun. 2009	Oct. 2009	5	Completed
Child 2	First spell	Sept.2008	Aug. 2009	12	Completed
Child 3	First spell	Sept. 2009	Feb. 2010	18	Right-censored

[8] Keep in mind that the first observed spell for a particular family may not be their actual first spell of participation. All we know is how many spells are observed in the study period. Those who have multiple spells are likely to be different from those with one spell in ways that may not be captured in the variables in the data.

[9] By definition, an analysis of completed spells would exclude all left-censored and all right-censored spells.

3.5 Summary

Although alternative approaches are sometimes used, the recommended approach to dealing with the issues of censoring and sampling in studies of spell length is to include the first observed non-left-censored spell per child (or family). This entry cohort approach avoids the oversampling of the (potentially) longer left-censored spells. By including right-censored spells, this approach avoids the bias of using only completed spells. Finally, by using only one spell per child (or family), this approach avoids weighting more heavily the experience of those with multiple spells (which may on average be shorter than spells for those with just one spell).
In summary, for each child (or family, if doing family-level analysis), the first observed spell that begins after the start of the study observation period (i.e., that is not left censored) is included in the study.

4. Appropriate Analytical Methods

There are a wide variety of analytical methods for event history analysis, and we cannot cover even a fraction of them in this brief. Instead we will focus on a few of the most commonly used methods, discuss their uses, and provide references for interested readers. [10]

Most analyses of program participation include estimates of median spell length and survival functions. Survival functions provide estimates of probability of surviving until a certain time.[11] In the case of subsidy participants, the survival function tells us the proportion who have not yet left the program as of a given time. Survival analysis methods account for the right censoring of spells that continue past the end of the observation window. Median spell length estimates should be reported rather than means because estimates of the mean are biased when some observations are censored. Also, given that there are likely to be a few very long spell lengths, it is preferable to report the median rather than the mean if the distribution of spell lengths is skewed.

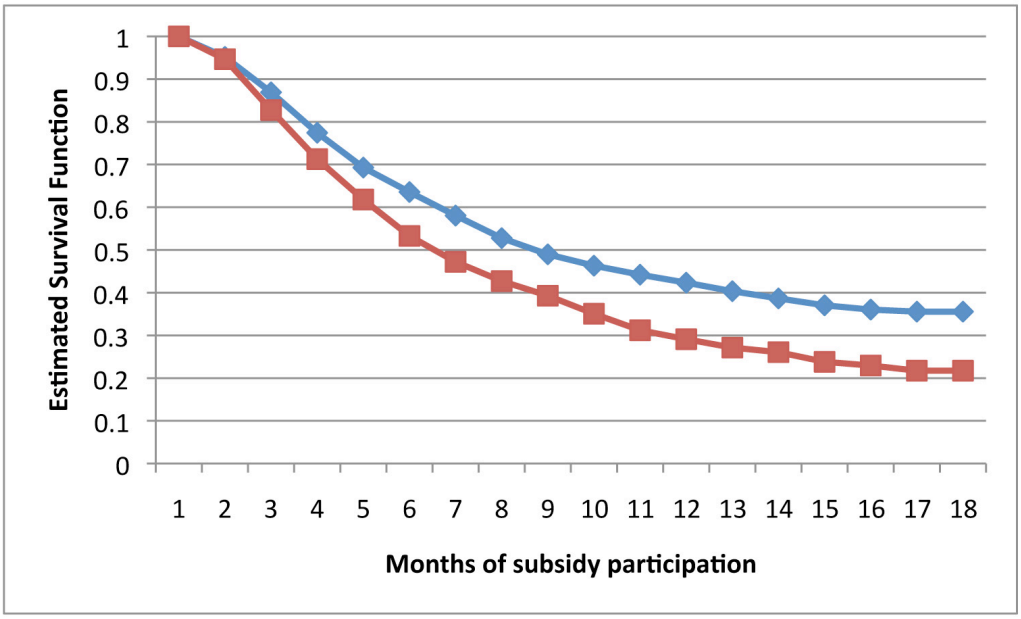
There are a number of different statistical techniques that can be used to estimate the distribution of spell lengths and the survival function. Which method is best depends on the type of data (e.g., continuous versus discrete time) and the type of model and estimation technique chosen (e.g., parametric or nonparametric). The references identified in Table 4 provide a good background on these options along with details on how to implement these methods in the commonly used statistical packages, SAS and STATA.

There are two basic methods of estimating spell lengths. The first is a nonparametric approach, Kaplan-Meier estimation. The second approach, commonly referred to as Accelerated Failure Time (AFT) models, is a parametric approach, meaning that the researcher makes assumptions about the functional form of the survival (or hazard) function. The most commonly reported statistics from these analyses are (i) the median spell length and (ii) the survival function. The median spell length tells us the length of time by which half of all families have exited the subsidy program. Other percentiles of the spell distribution, such as the 25th and 75th percentiles can also be estimated. These provide information on the length of short and long spells within the study population.

[10] Many general texts describe the analysis methods appropriate for survival analysis or event history analysis. Singer and Willett (2003) is thorough and accessible.
[11] The terminology reflects the fact that survival analysis methods were first developed primarily by biostatisticians studying death rates (Allison, 2011).

Figure 3 illustrates estimated survival functions, showing the proportion who have not yet left the program as of a given time for two subgroups. The survival functions start at 1.0 (no one has exited the program at time zero) and decline steadily. If one group has a survival curve always below another's, the first group exits more quickly and therefore has shorter spells. Statistical tests can be conducted to determine if the survival curves of two groups differ in a statistically significant sense.

Figure 3: Example of Estimated Survival Functions for Two Groups



Often we are interested in explaining differences in spell length across families or states. Multivariate models can be used to determine the factors that are associated with exit from the subsidy program. Rather than directly modeling the spell length as the dependent variable, these models use a binary dependent variable that indicates if each (monthly) observation is an exit month.[12] The probability of exit from the subsidy program is modeled as a function of child, family, and provider characteristics, as well as community or policy variables.[13] These models are typically estimated using either a proportional hazards model (also called Cox regression), or a parametric AFT model.[14] As with the estimation of median spell length, the decision about which model and estimation technique to use depends on the data and assumptions the researcher wants to make.

Other measures of subsidy continuity and duration include the number of spells and the cumulative number of months of subsidy participation within a fixed time period. One can also estimate the proportion who return to subsidy within three, six, nine or twelve months. If the focus of the study is on returns to the subsidy program, one would need to include multiple spells in the data set. One could measure the length of time between subsidy spells, and similar methods could be used to estimate median length of time from subsidy exit until return.

[12] The problem with using spell length as the dependent variable is that the distribution of spell lengths is highly likely to violate the assumption of normality needed for ordinary least squares regression (Cleves, et al. 2008).
[13] These characteristics are typically measured based on the first month of the spell or they may vary over time.
[14] Other methods can be used for these models, such as piecewise exponential models, logistic regression, and competing risk models. See Allison (2011) or Singer and Willett (2003) for more details.

5. Conclusion

Research over that past decade on the duration of participation in the child care subsidy program often produced results that were not comparable across studies because of the use of different methods. Although consensus on how to measure duration has emerged in related fields, no similar agreement exists in the field of child care and early education. Given the importance of these questions to families and policymakers, it is critically important that funders of the research have a basic understanding of the issues and methods of survival analysis. Studies that are based on designs using point-in-time samples or completed spells will produce results that are not comparable to other studies, and which are potentially misleading in their representation of the experience of families in the subsidy program. Other issues, such as whether to define spells based on one or two-month breaks, and whether to use child or family-defined spells, are important areas for future research. We encourage researchers to investigate and report how their results vary with these decisions so that over time recommendations may emerge on how to address them. Reliable and comparable findings on subsidy participation dynamics can assist both federal and state policymakers in designing child care subsidy programs.

Table 4. Analytic Methods Commonly Used in Studies of Dynamics of Subsidy Participation

Research Objective	Measures	Methods	Description	Examples of Statistical Software Commands	References
How long do spells of subsidy receipt last? Do spell lengths vary by subgroup?	Survival function (proportion remaining on program at different points in time); Distribution of subsidy spell lengths (median and other percentiles of spell lengths)	Kaplan-Meier estimation	+ Non-parametric method + Accounts for censoring of spells that continue past the observation window.	SAS lifetest	Allison (2011)
				STATA stset, sts list, sts graph	Cleves, Gould & Gutierrez (2008)
		Accelerated Failure Time (AFT) models	+ Assumes a particular functional form for the survival (or hazard) function + Accounts for censoring of spells that continue past the observation window.	SAS lifereg	Allison (2011)
				STATA streg	Cleves, Gould & Gutierrez (2008)
What factors are associated with exits from the subsidy program (and therefore with length of subsidy spells)?	Model the factors associated with the probability of exit from the subsidy program (results indicate the association between each factor and the probability of exit)	Proportional Hazard or Cox Regression model	Semi-parametric models of the characteristics associated with exiting the subsidy program, conditional on participating in the program.	SAS phreg	Allison (2011)
				STATA stcox	Cleves, Gould & Gutierrez (2008)
		Accelerated Failure Time (AFT) models	Parametric models of the characteristics associated with exiting the subsidy program, conditional on participating in the program.	SAS lifereg	Allison (2011)
				STATA streg	Cleves, Gould & Gutierrez (2008)

References and Related Works

Allison, P.D. (2011). *Survival Analysis Using the SAS® System: A Practical Guide*, 2nd ed. Cary, NC: The SAS Institute, Inc.

Bane, Mary Jo and David T. Ellwood (1986), "Slipping Into and Out of Poverty: The Dynamics of Spells" *Journal of Human Resources* 21(1), pp. 1-13.

Bane, Mary Jo and David T. Ellwood (1994), *Welfare Realities: From Rhetoric to Reform* (Cambridge, MA: Harvard University Press).

Berger, Mark C. and Black, Daniel A. (1998) "The Duration of Medicaid Spells: An Analysis Using Flow and Stock Samples" *Review of Economics and Statistics*, 80 (4), pp.667-675.

Cleves, Mario A., William W. Gould, and Roberto G. Gutierrez. 2008. *An Introduction to Survival Analysis Using Stata*, 2nd ed. College Station, TX: Stata Press.

Grobe, Deana, Roberta B. Weber and Elizabeth E. Davis. (2003). Guidebook for Implementing a Study on the Dynamics of Child Care Subsidy Use. Corvallis, OR: Oregon State University Family Policy Program, Oregon Child Care Research Partnership. March.

Ha, Yoonsook. (2009). Stability of Child-Care Subsidy Use and Earnings of Low-Income Families. *Social Service Review*, December: 495-523.

Herbst, Chris and Anna Johnson (2010). "Can We Trust Parental Reports of Child Care Subsidy Receipt?" Working paper.

Klerman, Jacob A. and Steven J. Haider (2001), "A Stock-Flow Analysis of the Welfare Caseload: Insights from California Economic Conditions" (Santa Monica, CA: Rand Corporation Labor and Population Program Working Paper Series 01-02).

Meyers, M. K., L. R. Peck, E. E. Davis, A. Collins, J. L. Kreader, A. Georges, R. Weber, D.T. Schexnayder, D. G. Schroeder, and J. A. Olson (2002), The Dynamics of Child Care Subsidy Use: A Collaborative Study of Five States (New York, NY: National Center for Children in Poverty).

Singer, Judith D. and John B. Willet. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford: Oxford University Press.

Swartz, Katherine, John Marcotte and Timothy D. McBride. (1993). Spells Without Health Insurance: The Distribution of Durations When Left-Censored Spells Are Included. *Inquiry* 30 (1, Spring): 77-83.

Weber, Roberta B. (2005). *Measurement of child care arrangement stability: A review and case study using Oregon child care subsidy data* (Doctoral dissertation, Oregon State University).

Witte, A. D. and M. Queralt. (2005). An Examination of the Duration of Child Care Subsidies in Rhode Island: Impacts of Policy Changes and Cross State Comparisons. Rhode Island: Wellesley College.

Wulczyn, Fred (1996), "A Statistical and Methodological Framework for Analyzing the Foster Care Experiences of Children" *Social Service Review* 70(2), pp. 318-329.

Wulczyn, Fred, Joseph Kogan and John Dilts (2001), "The Effects of Population Dynamics on Performance Measurement" *Social Service Review* 75(2), pp. 292-317.

Wulczyn, F., Joseph Kogan, and Brenda J. Harden (2003), "Placement stability and movement trajectories." *Social Service Review* 77(2), pp. 212-236.