

Evaluation Review

<http://erx.sagepub.com>

Designing a Household Survey to Address Seasonality in Child Care Arrangements

Stefanie R. Schmidt, Kevin H. Wang and Freya L. Sonenstein
Eval Rev 2008; 32; 216 originally published online Jan 11, 2008;
DOI: 10.1177/0193841X07311993

The online version of this article can be found at:
<http://erx.sagepub.com/cgi/content/abstract/32/2/216>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Evaluation Review* can be found at:

Email Alerts: <http://erx.sagepub.com/cgi/alerts>

Subscriptions: <http://erx.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 6 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://erx.sagepub.com/cgi/content/refs/32/2/216>

Designing a Household Survey to Address Seasonality in Child Care Arrangements

Stefanie R. Schmidt

U.S. Department of Education

Kevin H. Wang

RTI International

Freya L. Sonenstein

Johns Hopkins University Bloomberg School of Public Health

In household telephone surveys, a long field period may be required to maximize the response rate and achieve adequate sample sizes. However, long field periods can be problematic when measures of seasonally affected behavior are sought. Surveys of child care use are one example because child care arrangements vary by season. Options include varying the questions posed about school-year and summer arrangements or posing retrospective questions about child care use for the school year only. This article evaluates the bias associated with the use of retrospective questions about school-year child care arrangements in the 1999 National Survey of America's Families. The authors find little evidence of bias and hence recommend that future surveys use the retrospective approach.

Keywords: *child care; seasonality; recall bias; telephone survey*

Surveys that attempt to measure the activities of children and youth need to address seasonal variation that can occur in these activities. In the United States, for example, surveys that are conducted in the summer are likely to get very different estimates of how children spend their time compared to surveys that are conducted during the other seasons when school is in session. Ideally, the field period of a survey should correspond with the

Authors' Note: Please address correspondence to Stefanie R. Schmidt, Institute of Education Sciences, U.S. Department of Education, 555 New Jersey Avenue, NW, Washington, DC 20208; *phone:* (202) 219-2129; *fax:* (202) 219-1725; *e-mail:* Stefanie.Schmidt@ed.gov.

desired reference period for the questions of interest. However, a long field period that bridges the seasons may be required to maximize the response rate and achieve an adequate sample size.

Most household surveys that have measured child care have avoided seasonality issues by limiting the field period to months when school is in session. Those surveys include the 1990 National Child Care Survey, the Survey of Income and Program Participation, the Current Population Survey, and the National Household Education Survey. Two surveys, the National Survey of Families and Households (NSFH) and the National Longitudinal Survey of Youth 1979 (NLS-Y79), have fielded surveys that extended into the summer. The NSFH did not ask questions about child care arrangements of school-age children, for whom school-year and summer child care arrangements were most likely to differ. Analyses of the 1988 NLS-Y survey showed large differences in arrangements between those interviewed in summer and non-summer months for school-age children (Raley, Harris, and Rindfuss 2000). No survey has been designed to accurately measure summer child care arrangements since then.

Obtaining accurate estimates of the types of child care that are used by families throughout the year is important for the formulation of policy at the state and national levels as well as for predicting the implications of policy changes. Although the choice of what child care arrangements to use is primarily made by parents, aggregate utilization patterns can influence and are influenced by the substantial federal investments in child care resources. The federal programs include the Dependent Care Credit in the Internal Revenue Code, the Child Care and Development Grants, Temporary Assistance for Needy Families grants, the Social Services Grants, the Head Start Program under the Administration of Children and Families in the Department of Health and Human Services, and the Child Care Food Program under the U.S. Department of Agriculture. In fiscal year 2006, it has been estimated that federal investments included \$6.8 billion for Head Start, \$2.2 billion for the Child Care and Development Block Grant, and \$3 billion for child care entitlements to states (Steuerle, Reynolds, and Carasso 2007). The \$2.1 billion U.S. Department of Agriculture Child and Adult Care food program serves 1.8 billion nutritious meals and snacks per year to children ages 0 to 12 in afterschool programs, child care centers, and family day care homes serving primarily low-income children (Food Research and Action Center 2007). These programs each have different laws and regulations that often specify the types of arrangements that are eligible for subsidy. These programs can influence the types of arrangements parents choose, and

reciprocally, their expenditure levels can be influenced by the choices parents make. Hence, it is important to track the types of child care choices parents, especially low-income parents, are making and to consider the implications of seasonal variations in these choices, a policy question that has not been well examined heretofore.

In addition, many policy makers and researchers are interested in how many school-age children spend time in self-care and afterschool programs. Motivated by a concern that too many children are home unsupervised (potentially engaged in high-risk activities) and the belief that afterschool programs could provide a safe place for children to participate in academically or socially enriching activities (Hollister 2003), funding for afterschool programs has grown in recent decades. For example, the U.S. Department of Education's 21st Century Learning Centers grant program grew from \$40 million in 1998 to \$981 million in 2006 (Hollister 2003, U.S. Department of Education 2007). However, research on the detrimental effects of self-care is mixed and suggests that self-care may be harmful for only some children (Aizer 2004; Vandell and Shumow 1999). In addition, afterschool care programs do not necessarily reduce self-care, afterschool programs often do not improve academic outcomes, and afterschool programs sometimes increase negative behaviors (Dynarski et al. 2004; Kane 2004). Because attendance in afterschool care programs is often sporadic (Dynarski et al. 2004; Kane 2004), household survey data on the number of hours attended per week are particularly important to understand the extent to which children participate in such programs.

This article describes and evaluates the design of the child care component of the 1999 National Survey of America's Families (NSAF) to provide estimates of child care for the school year in which a significant portion of the sample was interviewed during the summer months when school was not in session. The 1999 NSAF was a primarily random-digit-dial (RDD) telephone survey of more than 42,360 households (with and without children) across the United States designed to produce both national- and state-level estimates for children and adults under age 65 for a broad range of measures covering health care, child care, education, program participation, and other measures of well-being. The 1999 survey was the second of three cross-sectional surveys conducted by Westat for the Urban Institute's Assessing New Federalism Project. For the 1999 survey, the overall child-level response rate for the combined telephone and nontelephone sample was 62.4% (Abi-Habib, Safir, and Triplett 2004).

Designing the 1999 Survey

The field period for the 1997 NSAF lasted from February to November. The length of the field period posed difficulties for obtaining estimates of child care during the school year for three reasons. First, the length of the field period introduced seasonal differences in estimates because of differences in summer and school-year child care arrangements. Second, those interviewed during the summer months were asked a different version of the child care module that reflected the diversity of child care arrangements during the summer, but these items were not truly comparable with items asked for the rest of the sample.

Finally, the respondents who were interviewed during the summer months were not a random sample of all respondents. Those interviewed during the summer months were likely to have required greater effort to interview (e.g., requiring more calls to complete the interview or more likely to have refused at some point to do the survey) and as such may have differed from those interviewed earlier on the key variables of interest (Groves and Wissoker 1999). As a result, our published child care estimates based on the 1997 data used only data obtained in the nonsummer months weighted to provide representative estimates of school-year child care. The estimates used a specially created set of weights that treated those interviewed during the summer months as nonrespondents. Anticipating that the field period for the 1999 NSAF would extend into the summer months, we altered the survey procedures so that we could systematically examine differences in child care between the summer and nonsummer months for a portion of the sample as well as obtain reliable estimates of care in the nonsummer months for the bulk of the sample.

The field period for the 1999 NSAF ran from February 15, 1999, to October 3, 1999. Respondents who were interviewed before June 13 and after September 26 were asked about their child care patterns during the last month. Those who were interviewed between June 13 and September 26, about 20% of all respondents, were asked retrospective questions about their child care arrangements during the month of May (Vaden-Kiernan, Cunningham, and Warren 2000). We chose May as a reference period because it was the last full month of the school year, and estimates for child care during the school year were desired.¹

When respondents are asked retrospective questions, the quality of survey data will depend on the respondents' ability to accurately recall the information required to answer the questions. In general, survey methodologists have found that the ability to recall events varies with the salience of events, the

degree to which the information required has been accessed in the past and most relevant for our study, and the length of time between the survey administration and the event being recalled (Eisenhower, Mathiowetz, and Morganstein 1991). The longer the period of time between the time of the interview and the event being recalled, the greater the difficulty in accessing the information required to answer the survey question. A longer length of time to recall does not in itself cause the memory of the event in question to fade. Instead, a longer recall period makes it more likely that the respondent will experience events similar to the initial event being recalled and that these additional events make recall of the initial event more difficult (Tourangeau, Rips, and Rasinski 2000). Previous research has not examined whether there is recall bias in retrospective questions about child care arrangements.

Cognitive testing of the 1999 NSAF survey instrument at the pilot phase showed that respondents were able to accurately recall May child care arrangements during interviews conducted in the month of August. There were two likely reasons for their lack of difficulty with the task. First, we were asking about a fairly recent event, as it was at most 4 months in the past. Second, for respondents with school-age children, the end of the school year seemed to be a salient event because they had to adjust child care arrangements. For parents of children in preschool, the end of the school year can also be a salient event because many preschools are closed during the summer (Capizzano, Adelman, and Stagner 2002).

Whether a respondent was administered the May retrospective version of the questionnaire was not random. As in many RDD surveys, sampled telephone numbers were released in replicates called release groups. Interviewers began working with almost all release groups well before June, so the May retrospective respondents were among the most likely to be difficult to contact and/or interview. Responses from the May retrospective respondents are combined with the respondents who were interviewed during the school year, to obtain overall school-year estimates.

Estimates of summer child care patterns were produced using data collected from the summer release group, the final release group in the 1999 NSAF. The summer release group was released on June 13, and the vast majority of the interviews (99% of cases) were completed during the summer. The respondents in the summer release group answered a version of the questionnaire that asked about child care during the last month and included a larger set of questions to capture the diversity of arrangements in the summer. The summer release group composed 4% of the overall sample.

Using data from the 1999 NSAF, Capizzano et al. (2002) document some important seasonal differences in school-age child care. In particular, during

the summer months, children spend time in summer school and summer programs, arrangements not available during the school year. In addition, although children ages 6 to 12 with employed parents are no more likely to be spending time in other types of supervised child care arrangements (such as relative care, nanny/babysitter care, or family child care) during the summer than during the school year, the number of hours per week spent in those arrangements increases from 13.9 hr during the school year to 23.2 hr during the summer months. Similarly, school-age children are no more likely to be in self-care during the summer months than during the school year, but the number of hours per week increases from 4.8 hr per week during the school year to 10.3 hr per week during the summer months.²

In addition, preschool children also have some important seasonal differences in their child care arrangements. Three- and 4-year-olds are much less likely to be in center-based care during the summer months; 50% are in center-based care during the school year compared with 28% during the summer months.

Raley et al. (2000) examined seasonal differences in child care arrangements using the 1988 interview data from NLS-Y 79 Child-Mother Sample. They found large seasonal differences in child care for school-age children but, unlike Capizzano et al. (2002), did not find large seasonal differences for preschool children. There are several possible reasons for the differences in findings between the two studies: the 11-year gap in when the surveys were conducted, differences between the two surveys in the ways the questions about child care were asked, and the fact that Raley et al. (2000) estimates did not take into account the nonrandom differences in interview timing in the NLS-Y 79. Like the 1999 NSAF, the 1988 wave of the NLS-Y had a long field period (the 1988 interviews took place between June and December, with 82% of interviews taking place during the summer months), and refusal conversion efforts probably meant that more reluctant respondents were more likely to be interviewed in the later months of the survey during the school year (Center for Human Resource Research 2006).

Method

We examined overall school-year estimates for two types of biases that may arise because of the use of retrospective questions for a significant portion of the sample. First, we examined the possibility of recall bias among our respondents who were asked retrospective questions about their child care during the month of May. We test for recall bias by comparing responses

obtained from interviews in which respondents were asked about child care during May retrospectively with those who were asked about child care during May in the standard interview for the survey. Second, we examine our overall school-year child care estimates for bias because of the fact that a disproportionate share of our sample was asked questions about child care during May. If there are large differences between May child care and child care used during the rest of the school year (between February-April and September-October), our overall estimates of school-year child care arrangements would be biased because a disproportionate share (37%) of the school-year interviews were conducted using data from May child care arrangements.

In the NSAF, child care arrangements are reported for up to two randomly selected children (or focal children) per household, one selected from all children under the age of 6 and one selected from all children 6 to 12 years old. Our analyses of arrangements differentiate between those made for preschool-age children 0 to 4 years old and those made for school-age children 6 to 12 years old. Within each of these comparisons, we examine the use of all child care arrangements in which a child was placed. We examine whether the child was in a particular type of arrangement on a regular basis in the last month and, conditional on being in that type of arrangement, how many hours per week the child spent in that arrangement. Our child care arrangement categories are not mutually exclusive because many children were regularly in more than one arrangement. Unlike some other analyses of NSAF data (Capizzano et al. 2002; Sonenstein et al. 2002), we are not looking at the main or primary child care arrangements, and we are reporting child care used by children of both employed and nonemployed respondents in our sample. In addition, we also count school as a child care arrangement.³

Our analyses focus on the telephone (RDD) portion of the 1999 NSAF sample. We excluded the nontelephone sample from our analyses because these cases were not included among the telephone release groups. The NSAF interview process consisted of a screener interview to determine household eligibility followed by extended interviews with sampled household members to collect substantive information. Thus, the overall response rate for child estimates is the product of the screener and extended interview response rates. For the 1999 NSAF, the child response rate for the RDD sample was 61.4% (Abi-Habib et al. 2004).

Although the NSAF response rate is quite respectable compared to most RDD surveys, estimates from the survey are still susceptible to bias because of nonresponse. To adjust for this potential source of bias, our analyses used survey weights that were created to compensate for differences between respondents and nonrespondents. The NSAF weights include adjustments for

nonresponse in which the weights of respondents were adjusted to account for the weights of nonrespondents using variables available for both respondents and nonrespondents. In addition, the nonresponse adjusted weights were further adjusted to match population control totals from the U.S. Census Bureau on age, race/ethnicity, gender, and home ownership (Brick et al. 2000).⁴

Models were estimated with SUDAAN Version 8 (RTI International 2001). Standard errors were estimated by using the NSAF focal child replicate weights, which use a paired jackknife approach to account for the complex survey design features of the NSAF.

Given that we conduct several tests for differences of proportions and means for use and hours of care, it is possible that some of the estimated bias coefficients will be statistically significant, even if bias does not exist. For instance, when assessing statistical significance at the 5% level, we would expect that approximately 5% of the tests for differences in means would be statistically significant, even if there were no true bias, simply because of chance alone (Benjamini and Hochberg 1995; Jackson et al. 2007). Because of this potential bias, we use the Benjamini-Hochberg approach for adjusting for multiple comparisons, which adjusts the critical p values for statistical significance.⁵ Because some readers will be interested in statistical tests for individual variables such as self-care rather than an entire domain, we also note whether individual t tests that were statistically significant at the .05 significance level unadjusted for multiple comparisons. In Tables 1 and 2, we make adjustments for multiple comparisons with each of four domains: percentage children in arrangements for non-May school-year interviews, percentage of children in arrangement for May retrospective, average hours of care in non-May school-year interviews, and average hours of care for May retrospective.

Results

We test for recall bias by comparing child care arrangements reported by respondents who were interviewed during the summer (after June 13) and asked about May child care, with the child care arrangements reported by those who were interviewed between May 13 and June 13 about child care used during the previous 4 weeks. We refer to these two samples as the "May retrospective" and "May current" samples, respectively. We also test for differences in child care arrangements in May and the rest of the school-year interview period (February to April and late September to October). We compare the answers of school-year respondents who were interviewed about

Table 1
School-Year Child Care Arrangements of Children
Ages 6 to 12 by Type of May Interview

	May Current		May Retrospective		Non-May School-Year Interviews	
	Percentage Children	Average Hours	Percentage Children	Average Hours	Percentage Children	Average Hours
School (regular or summer)	99	33	96	33	100	33
Any nonparental child care other than school	52	11	49	12	51	12
Before- and/or afterschool care	15	11	15	11	18	11
Family child care	8	9	7	12	8	10
Relative	26	12	27	14	24	12
Nanny/babysitter	6	9	5	9	7	8
Self-care	16	4	13 ^a	5	15	4
Sample size	2,042		2,619		7,995	

a. Difference when compared to May current sample is statistically significant at the .05 level.

child care used during months other than May (those with interviews between February and mid-May or late-September and October) with answers of the current May respondents. We will refer to the first sample as the “non-May school-year sample.”⁶ Estimates of the percentage of children reported in each type of child care arrangement in each of the three samples, along with average hours reported in each arrangement, are shown in Table 1 for children 6 to 12 years old and Table 2 for children under 5 years old.

Our descriptive analysis points mostly to a lack of recall bias in child care arrangements of 6- to 12-year-olds, although there is weak evidence of recall bias for one of the 14 variables prior to adjustments for multiple comparisons. Parents of children in the May retrospective sample were less likely to have reported self-care (13%) than parents of children in the May current sample (16%), but the difference was not statistically significant after the adjustment for multiple comparisons was made. In the other 13 variables measuring use of or hours in a child care arrangement for school-age children, we do not find any evidence of recall bias. Our descriptive analyses find no evidence of recall bias in child care arrangements of 0- to 4-year-olds.

Turning to our tests for possible biases because of differences between non-May school-year reported child care arrangements and current May child care arrangements, we find no statistically significant differences between

Table 2
School-Year Child Care Arrangements of Children
Ages 0 to 4 by Type of May Interview

	May Current Sample		May Retrospective Sample		Non-May School-Year Interviews	
	Percent Children	Average Hours	Percent Children	Average Hours	Percent Children	Average Hours
Any nonparental child care	68	26	63	26	68	26
Center-based care	30	26	27	24	30	25
Family child care	15	23	12	25	14	24
Relative	33	20	33	21	35	20
Nanny/babysitter	8	16	7	14	7	14
Sample size	1,519		1,942		6,146	

reported child care arrangements in May and the rest of the school year for 6- to 12-year-olds (Table 1) or 0- to 4-year-olds (Table 2).

Additional testing for bias was carried out by estimating two types of regression models—logistic regression models with the dependent variable measuring whether a child was in a particular kind of arrangement and ordinary least squares (OLS) models with the dependent variables measuring the number of hours a child was in an arrangement, conditional on using that type of arrangement. For school-age children, models with school as the dependent variable were not estimated because of the lack of variation in the dependent variable—nearly all children were in school. For each model estimated, we pooled all respondents who were interviewed about school-year child care and included two categorical independent variables indicating whether the respondent was in the May retrospective sample or the non-May school-year sample (the reference group is the May current sample). A statistically significant coefficient on the May retrospective variable would be an indicator of recall bias, demonstrating that parents interviewed during the summer months and asked retrospective questions about May child care arrangements answered differently than parents interviewed in late May and early June about child care arrangements used during the past month. A statistically significant coefficient on the non-May school-year variable would be an indicator that child care arrangements used during the month of May differ from child care arrangements used during the remainder of the school year.

Each model included additional variables to control for possible differences between the samples. Several variables measure the respondents' demographic and labor market characteristics, including categorical variables for whether the respondent was married, female, employed, working full-time, holding more than one job, self-employed, a current welfare recipient, ever received welfare, had attended some college, and a college graduate and had work hours falling between 6 a.m. and 6 p.m. and a continuous variable measuring the respondent's monthly earnings. Other variables measure the respondent's spouse or partner's labor market status, including categorical variables for whether he or she was employed, was working full-time, had work hours falling between 6 a.m. and 6 p.m., held more than one job, or was self-employed. For the focal child, categorical variables are included to measure the child's age and race/ethnicity (non-Hispanic Black or Hispanic). In addition, variables are included to measure the total number of children in the family ages 0 to 5, the total number of children ages 6 to 12, state of residence, and an indicator variable for whether the family's income was less than or equal to 200% of the federal poverty level. Many of these variables are standard in the child care literature, and most have been used in other multivariate analyses of child care arrangements using the NSAF (Tout et al. 2003). Finally, we included an indicator of whether the respondent had ever refused to do the survey as an indicator of interviewing difficulty. We found that 61.5% of the May retrospective respondents had ever refused to participate in the NSAF as compared with 55.6% of May current and 23.8% of non-May school-year respondents.⁷

For the logistic regression models, we report the predictive margins for the May current, May retrospective, and non-May school-year variables. The predictive margin can be interpreted as the average predicted percentage in that care arrangement if all individuals in the population had been in that group, either the May current, May retrospective, or non-May school-year group (Korn and Graubard 1999). We also provide differences between the predictive margins for the May current and May retrospective groups and the May current and non-May school-year groups. For the OLS regression models, we provide the regression coefficients as well as predictive margins.

As with the descriptive analyses, the multivariate analyses point away from recall bias. Only two variables, use of self-care and hours in family child care, showed evidence of recall bias prior to adjustments for multiple comparisons, but those differences did not remain after the adjustments for multiple comparisons are made. Retrospective May respondents were 3.2 percentage points less likely to report self-care among 6- to 12-year-olds than current May respondents, which was consistent with the descriptive analyses,

Table 3
Results From Logistic Regression Models Predicting Use of Child Care Arrangements for Children Ages 6 to 12

	Predictive Margin	Standard Error	Difference From May Current	Standard Error of Difference	Pseudo R^2
Any nonparental care					.1590
May current	.4796	.0198			
May retrospective	.4730	.0168	.0066	.0243	
Non-May school	.4836	.0098	-.0041	.0228	
Before- or afterschool care					.1010
May current	.1538	.0145			
May retrospective	.1563	.0124	-.0024	.0179	
Non-May school	.1778	.0070	-.0240	.0169	
Family care					.0510
May current	.0820	.0103			
May retrospective	.0790	.0093	.0030	.0139	
Non-May school	.0803	.0048	.0017	.0113	
Relative care					.0639
May current	.2614	.0136			
May retrospective	.2576	.0138	.0038	.0183	
Non-May school	.2448	.0081	.0166	.0160	
Nanny/babysitter care					.0379
May current	.0614	.0089			
May retrospective	.0583	.0084	.0031	.0131	
Non-May school	.0694	.0055	-.0080	.0103	
Self-care					.1181
May current	.1625	.0146			
May retrospective	.1292 ^a	.0106	.0333 ^a	.0156	
Non-May school	.1515	.0062	.0110	.0158	

a. Difference when compared to May current sample is statistically significant at the .05 level.

but the difference was not statistically significant when adjustments for multiple comparisons were made (see Table 3). In addition, retrospective May respondents reported an average of 3.6 fewer hours per week in family child care than current May respondents, a difference that did not remain when adjustments for multiple comparisons were made (see Table 4).

Like the descriptive results, the multivariate results showed no evidence of statistically significant differences in current May and non-May school-year child care arrangements for children ages 6 to 12. For children ages 0 to 4, the multivariate analysis also showed no evidence of statistically significant differences in survey estimates (for both use of arrangements and hours in arrangements) between the current May, retrospective May, and non-May school-year respondents.⁸

Table 4
Results From OLS Regression Models Predicting Hours in Child Care Arrangements for Children Ages 6 to 12

	Estimate	Standard Error	Predictive Margin	Standard Error of Predictive Margin	R ²
Any nonparental care					0.1357
May current			13.0357	0.6552	
May retrospective	0.8282	0.8451	13.8639	0.5579	
Non-May school	-0.2098	0.6455	12.8259	0.3529	
Before- and/or afterschool care					0.1278
May current			11.3111	0.7508	
May retrospective	-0.6459	0.8440	10.6652	0.6318	
Non-May school	-0.3706	0.8148	10.9404	0.4046	
Family care					0.2298
May current			9.5715	0.6832	
May retrospective	3.4061 ^a	1.5606	12.9776 ^a	1.4886	
Non-May school	0.4017	0.8381	9.9732	0.4923	
Relative care					0.1352
May current			12.0281	0.9722	
May retrospective	1.4039	1.1927	13.4320	0.7838	
Non-May school	-0.2273	1.0507	11.8008	0.5456	
Nanny/babysitter care					0.3042
May current			8.3078	1.4826	
May retrospective	-0.6592	1.7929	7.6486	1.1064	
Non-May school	-0.0566	1.5397	8.2512	0.8935	
Self-care					0.0968
May current			4.3409	0.4408	
May retrospective	0.5923	0.6147	4.9332	0.3551	
Non-May school	-0.0534	0.4857	4.2875	0.2303	

a. Difference when compared to May current sample is statistically significant at the .05 level.

Readers particularly interested in use of self-care or hours in family child care may focus on the findings of bias prior to the adjustments for multiple comparisons and may be interested in understanding the magnitude of the bias. For the two variables in which we found statistically significant differences between the retrospective May and current May samples and percentage of children ages 6 to 12 in self-care and hours per week in family care, we reestimated additional models to simulate what our “synthetic school-year” point estimates of child care would have been in the absence of recall bias. Synthetic school-year estimates of child care come from pooling the retrospective May, current May, and non-May school-year samples. In these

models, we grouped the current May and non-May school-year samples together and used a single indicator for the retrospective May sample. The predictive margins from these regressions give us a sense of what our estimates might have looked like if we had not carried out any interviews during the summer. For the percentage of children ages 6 to 12 in self-care, the predictive margin was 15.4% (standard error of 0.58%) as compared to the actual estimate of 14.9% (standard error of 0.56%). For the average hours in family care for children ages 6 to 12, the predictive margin was 9.9 hr per week (standard error of 0.43 hr), which is close to the actual estimate of 10.5 hr per week (standard error of 0.43 hr). Thus, the use of retrospective questions for about 20% of the sample did produce a small bias in estimates of child care based on the synthetic school-year sample. Because the purpose of the May retrospective sample was to increase the precision of the estimates by increasing the sample size, one measure of the magnitude of the bias is to compare the bias with the size of the standard error of the point estimate. For both the use of self-care for school-age children and hours in family child care, the difference in the predictive margin was slightly higher than the standard error.

Conclusion

The results of this article should inform the design of future household surveys that measure child care arrangements and require interview dates spanning both the school year and summer. Because there is evidence of seasonal differences in child care arrangements and expenses, particularly among families with school-age children, future surveys of child care arrangements need to address this issue. We unequivocally recommend that if accurate estimates of child care arrangements during the summer are desired, the survey should include a separate summer replicate that asks questions about summer child care arrangements, as was done in the 1999 NSAF.

How should data from summer respondents not in the summer replicate be collected and used in estimates of school-year child care? One approach is to not ask any questions about child care during the summer months and treat summer respondents as nonrespondents and incorporate an additional nonresponse adjustment to adjust the weights of those interviewed during the school year to account for summer respondents. Estimates of school-year child care arrangements would use only data collected during the non-summer months.

There are several disadvantages to this approach, beginning with a loss of sample size and a corresponding reduction in precision of estimates.⁹ In addition, creating an additional weight for child care estimates may be confusing to users of the data. Finally, if child care questions are simply not asked of respondents being interviewed during the summer, there is a risk of affecting responses to other items in the survey by introducing the potential for context effects. If some respondents are asked about child care before other questions and other respondents are not asked these child care questions, responses to the subsequent questions may differ if the child care questions provide a different context for answering the subsequent questions.

We recommend the 1999 NSAF approach, which was to ask summer respondents retrospective questions about their May child care arrangements and to obtain estimates of school-year child care arrangements by combining data collected from the May retrospective respondents and data collected from the school-year interviews.

The disadvantage of the May retrospective approach is the potential for bias because of using a longer recall period for one group of respondents.¹⁰ If the field period cuts across the school year and summer, we recommend using the 1999 NSAF's approach for measuring child care arrangements of preschool children. The May retrospective approach does not appear to introduce statistically significant biases for estimates of 0- to 4-year-olds' school-year child care arrangements.

For asking survey questions about school-age children, we also recommend using the May retrospective approach. For 10 of our 12 child care arrangement variables, the use of the May retrospective approach produced no statistically significant differences in estimates, so we recommend the use of the May retrospective approach. In addition, when adjustments for multiple comparisons are made, we find no evidence of bias from the May retrospective approach for any of the child care arrangement variables. For analysts interested primarily in measuring self-care rather than the totality of school-age child care use, we recommend that a retrospective question approach be supplemented with other approaches for reducing recall bias. With respect to the questionnaire, there are various methods that can be used to improve recall, including the use of questions that help the respondent define the reference period and other related events during that period (Eisenhower et al. 1991; Tourangeau et al. 2000). Another approach would be to field a release group that would be asked questions about child care in May that could be used by analysts to detect any bias in self-care yet have the advantage of obtaining the additional sample size from the May retrospective sample.

This article could inform the design of future surveys that have long field periods and measure social indicators that have large seasonal differences between the summer and school year. Such indicators could include youth employment and school attendance. Designers of those surveys should consider including a separate summer replicate to accurately capture summer behavior and ask other summer respondents retrospective questions.

Notes

1. The NSAF RDD sample frame is supplemented with an area probability sample to conduct interviews with nontelephone households. We dropped the nontelephone sample from our analyses because the respondents were not included among the telephone release groups. In practice, estimates produced excluding the nontelephone sample differ only slightly from estimates produced including the nontelephone sample.

2. Capizzano et al. (2002) also found significant differences in summer child care arrangements for school-age children with nonemployed primary caregivers.

3. We do not include 5-year-olds in this analysis because they are difficult to group with either school-age or preschool children. Depending on school districts and birthdates, some 5-year-olds are already in school, while others have not yet started (Capizzano et al. 2002). Because of a limited sample size, we could not provide separate analyses restricted to 5-year-olds.

4. The response rate corresponds with Response Rate 4 as described by the American Association for Public Opinion Research (2006). The adjustment for unknown residential telephone status is based on the survival method described in Brick, Montaquila, and Scheuren (2002).

5. When only one t test is statistically significant at the conventional .05 level, the critical p value using the Benjamini-Hochberg approach is equivalent to the critical p value for the more conservative Bonferroni adjustment and is equal to $.05/n$, where n is equal to the number of t tests conducted in the domain.

6. We do not compare the non-May school-year sample with the retrospective May sample because it would be difficult to interpret whether any differences were due to recall bias or differences between May and non-May arrangements.

7. Full output for all regressions is available on request from the corresponding author.

8. Additional tables showing regression results for children ages 0 to 4 are available on request from the corresponding author.

9. In addition, an additional weighting adjustment for nonresponse will increase the variability of estimates if the variables used in the weighting are only related to survey nonresponse. Nonresponse biases will be reduced only if the variables used in the nonresponse adjustment are related to both the survey items of interest (in this case, child care measures) and survey nonresponse (Little and Vartivarian 2005).

10. We should note that in the case of the 1999 NSAF, because a summer replicate was used, an additional weight was needed to produce child care estimates for the school year anyway.

References

- Abi-Habib, N., A. Safir, and T. Triplett. 2004. *2002 NSAF survey methods and data reliability*, 2002 NSAF Methodological Report No. 1. Washington, DC: Urban Institute.
- Aizer, A. 2004. Home alone: Supervision after school and child behavior. *Journal of Public Economics* 88: 1835–48.
- American Association for Public Opinion Research. 2006. *Standard definitions: Final dispositions of case codes and outcome rates for surveys*, fourth edition. Lenexa, KS: Author.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling for the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1): 289–300.
- Brick, J. M., P. Broene, D. Ferraro, T. Hankins, and T. Strickler. 2000. *1999 NSAF sample estimation survey weights*, 1999 NSAF Methodological Report No. 3. Washington, DC: Urban Institute.
- Brick, J. M., J. Montaquila, and F. Scheuren. 2002. Estimating residency rates for undetermined telephone numbers. *Public Opinion Quarterly* 66 (1): 18–39.
- Capizzano, J., S. Adelman, and M. Stagner. 2002. *What happens when the school year is over? The use and costs of child care for school-age children during the summer months*. Washington, DC: Urban Institute.
- Center for Human Resource Research. (2006). *NLSY79 user's guide: A guide to the 1979–2004 National Longitudinal Survey of Youth Data*. Columbus: The Ohio State University. Available at www.nlsinfo.org/pub
- Dynarski, M., S. James-Burdumy, M. Moore, L. Rosenberg, J. Deke, W. Mansfield, and E. Warner. 2004. *When schools stay open late: The national evaluation of the 21st century community learning centers program new findings*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Eisenhower, D., N. Mathiowetz, and D. Morganstein. 1991. Recall error: Sources and bias reduction techniques. In *Measurement errors in surveys*, ed. P. P. Biemer, R. M. Groves, and L. E. Lyberg, pp. 127–144. New York: John Wiley and Sons.
- Food Research and Action Center. 2007. *Afterschool guide: Funding opportunities and nutrition resources for afterschool programs*. Washington, DC: Author.
- Groves, R. and D. Wissoker. 1999. *1997 NSAF early nonresponse studies*, 1997 NSAF Methodological Report No. 7. Washington, DC: Urban Institute.
- Hollister, R. 2003. *The growth in after-school programs and their impact*. Washington, DC: Brookings Institution.
- Jackson, R., A. McCoy, C. Pistorino, A. Wilkinson, J. Burghardt, M. Clark, C. Ross, P. Schochet, P. Swank, and S. Schmidt. 2007. *National evaluation of early reading first: Final report*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Kane, T. J. 2004. *The impact of after-school programs: Interpreting the results of four recent evaluations* (William T. Grant Foundation working paper). New York: William T. Grant Foundation.
- Korn, E. L., and B. I. Graubard. 1999. *Analysis of health surveys*. New York: Wiley and Sons.
- Little, R. J. A., and S. Vartivarian. 2005. Does weighting for nonresponse increase the variance of survey means? *Survey Methodology* 31(2): 161–68.
- Raley, R. K., K. M. Harris, and R. R. Rindfuss. 2000. Quality and comparability of child care data in U.S. surveys. *Social Science Research* 29: 356–81.

- RTI International. 2001. *SUDAAN user's manual: Release 8.0*. Research Triangle Park, NC: Author.
- Sonenstein, F. L., G. Gates, S. R. Schmidt, and N. Bolshun. 2002. *Primary child care arrangements of employed parents: Findings from the 1999 National Survey of America's Families*. Washington, DC: Urban Institute.
- Steuerle, C. E., G. Reynolds, and A. Carasso. 2007. *Investing in children*. Washington, DC: The Urban Institute.
- Tourangeau, R., L. Rips, and K. Rasinski. 2000. *The psychology of survey response*. Cambridge, UK: Cambridge University Press.
- Tout, K., M. Zaslow, C. Calkins, and J. Capizzano. 2003. *Unsupervised time: Family and child factors associated with self care*. Washington, DC: Urban Institute.
- U.S. Department of Education. 2007. *21st Century Learning Centers funding status*. Available at www.ed.gov/programs/21stcclcfunding.html
- Vaden-Kiernan, N., P. Cunningham, and P. Warren. 2000. *1999 NSAF telephone survey methods*, 1999 NSAF Methodological Report No. 9. Washington, DC: Urban Institute.
- Vandell, D. L., and L. Shumow. 1999. After school child care programs. *The Future of Children* 9(2): 64–80.

Stefanie R. Schmidt is an economist at the Institute of Education Sciences at the U.S. Department of Education. From 1998 to 2002, she was a research associate and then a consultant at the Urban Institute, where she worked primarily on child care research under the auspices of the Assessing the New Federalism Project.

Kevin H. Wang, PhD, is a survey methodologist in the Survey Research Division of RTI International. From 1996 to 2004, he was a research associate within the Urban Institute's Assessing the New Federalism Project, working on the National Survey of America's Families.

Freya L. Sonenstein is a professor at the Johns Hopkins Bloomberg School of Public Health. From 1995 to 2001, she directed the social services research agenda of the Assessing New Federalism Project at the Urban Institute.