
Appendix D. Data Quality

Imputation procedures. Two principal determinants of the quality of data collected in household surveys are the magnitude of the imputed responses and the accuracy of the responses that are provided. This appendix provides information on the imputation rates for selected child care items in the Survey of Income and Program Participation. The Fall 1991 data include interviews conducted from October 1991 to January 1992 and refer to child care arrangements used in the month prior to the survey.

Imputed responses refer either to missing responses for specific questions or "items" in the questionnaire or to responses that were rejected in the editing procedure because of improbable or inconsistent responses. An example of the latter is when a 14-year-old child is said to be cared for in a nursery school during the time his or her parent is at work.

The estimates shown in this report are produced after all items have been edited and imputed whenever necessary. Missing or inconsistent responses to specific items are assigned a value in the imputation phase of the data processing operation. The procedure used to assign or impute most responses for missing or inconsistent data for the SIPP is commonly referred to as the "hot deck" imputation method. The process assigns item values reported in the survey by respondents to nonrespondents. The respondent from whom the value is taken is called the "donor." Values from donors are assigned by controlling edited demographic and labor force data available for both donors and nonrespondents. The control variables used for child care items generally included the age of the child for whom there was missing data, the parent's marital status, and whether the parent was employed full- or part-time, looking for work or attending school.

Item nonresponses. Imputation rates for both primary and secondary child care arrangements (items 3a and 4a in the questionnaire shown in appendix E) for the respondents' three youngest children are shown in table D-1. The imputation rates are calculated by dividing the number of missing or inconsistent responses by the total number of responses that should have been provided based on the number of children in the household who required child care responses. In general, the level of imputation for primary child care arrangements for employed women in the SIPP panels in this report

averaged about 9 percent. This was up from a 7-percent imputation rate for the 1988 data panel. Lower imputation rates were found for secondary arrangements (about 3 percent).

Table D-2 shows imputation rates for other selected items. About 14 percent of the responses concerning whether a cash payment was made for the child's primary child care arrangement were imputed; another 7 percent failed to answer the question if any cash payment was made for secondary child care services. For those who were determined to have made a cash payment, about 17 percent failed to report on the amount of the payment for the primary arrangement while 11 percent failed to report the cash amount for the secondary arrangement.

Imputation rates for cash payment items were also higher in this survey than in previous years. In the fall 1988 survey, imputation rates of 10 percent and 4 percent, respectively, were recorded for failures in reporting on whether a cash payment was made for child care for either the primary or secondary arrangement. Imputation rates of 13 and 9 percent, respectively, were also reported for the amounts of the primary and secondary payments.

In cases where two or more children shared the same arrangement and when only one payment was made for the arrangement, respondents were asked to indicate which children shared arrangements and the total cost for the shared arrangement. Approximately 16 percent of the respondents failed to indicate if the primary arrangements were shared and another 10 percent failed to indicate if the secondary arrangements were shared. Hence, an additional degree of uncertainty was added to the procedure which ultimately derived the total cost of all arrangements.

Hours spent in child care. Approximately 16 to 17 percent of respondents in the survey had their responses imputed on the number of hours their children spent each week in child care. Hours that the child spent commuting to school or to the arrangement were not counted as part of the arrangement for several reasons. First, travel time on a bus is clearly not equivalent to time spent under a provider's supervision. Researchers attempting to estimate the time children spend in day care centers or nursery schools would not want to include supervision by a bus driver in their estimates.

Secondly, since child care costs per hour were computed in this report, adding unpaid travel time to the arrangement time would clearly bias the hourly child care costs downward.

The reader should also be aware that these estimates probably contain rounding errors resulting from the respondent mentally computing weekly estimates

Table D-1. Imputation Rates for Primary and Secondary Child Care Arrangements for Children Under 15 Years: Fall 1991

[Data represent actual numbers of arrangements mentioned in the survey. Data are shown for arrangements for all children under 15 years of age with parents in the labor force or in school]

Type of arrangement	Primary arrangement ¹			Secondary arrangement ²		
	Total	Imputed responses		Total	Imputed responses	
		Number	Percent		Number	Percent
Total	5,286	487	9.2	1,664	50	3.0
Child's other						
parent/stepparent ..	580	51	8.8	341	11	3.2
Child's brother/sister ..	54	2	3.7	120	9	7.5
Child's grandparent ..	360	45	12.5	295	5	1.7
Other relative of						
child.....	168	16	9.5	99	2	2.0
Nonrelative of child ..	457	41	9.0	289	10	3.5
Day/group care center	306	31	10.1	136	2	1.5
Nursery/preschool ...	133	6	4.5	34	-	-
School-based activity.	109	14	12.8	72	-	-
Kindergarten/grade school.....	2,827	262	9.3	60	1	1.7
Child cares for self...	95	11	11.6	179	8	4.5
Parent works at						
home.....	126	6	4.8	18	1	5.6
Parent cares for child at work	64	2	3.1	21	1	4.8
Child not born as of last month	7	-	-	-	-	-

- Represents zero.

¹Item 3a in child care questionnaire (appendix E).

²Item 4a in child care questionnaire (appendix E).

from daily time schedules which may involve fractional hours. The specificity of the question does not necessarily result in an equivalently accurate estimate. Estimating intervening travel between arrangements, which could involve several different trips over the course of a typical grade-school-age child's day, could involve memory and computational errors large enough to make these estimates less than reliable.

Table D-2. Imputation Rates for Selected Child Care Items: Fall 1991

[Data represent actual numbers of arrangements mentioned in survey. Data are shown for arrangements for all children under 15 years of age with parents in the labor force or in school]

Item number	Question	Number of arrangements		
		Total	Imputed responses	
			Number	Percent
3b	Place of arrangement			
4b	Primary arrangement	1,619	187	11.6
	Secondary arrangement ...	1,144	38	3.3
3c	Any payment made?			
4c	Primary arrangement	1,533	212	13.8
	Secondary arrangement ...	925	62	6.7
	Is payment shared?			
3d	Primary arrangement	652	107	16.4
4d	Secondary arrangement ...	361	37	10.2
	Amount of payment			
3e	Primary arrangement	989	168	17.0
4e	Secondary arrangement ...	493	56	11.4
	Weekly hours of care			
3f	Primary arrangement	5,279	816	15.5
4f	Secondary arrangement ...	1,664	287	17.2
	Changes in arrangements			
5a	For children under 5.....	1,625	194	11.9
5b	For children 5 to 14.....	3,654	635	17.4
5c	Number of changes	1,150	183	15.9

Note: Item numbers refer to child care questionnaire (appendix E).