APPENDIX A

ESTIMATES OF SAMPLING ERRORS

Two types of errors affect the estimates from a sample survey: (1) nonsampling errors and (2) sampling errors. Nonsampling errors are the result of errors committed during data collection and data processing, such as failure to locate and interview the correct household, misunderstanding of the questions on the part of either the interviewer or the respondent, and data entry errors. Although numerous efforts were made during the implementation of NFHS-2 to minimize nonsampling errors, they are impossible to avoid and difficult to evaluate statistically.

Sampling errors, on the other hand, can be evaluated statistically. The sample of women selected in NFHS-2 is only one of many samples that could have been selected from the same population, using the same design and expected sample size. Each of these samples would yield results that differ somewhat from the results of the actual sample selected. The sampling error is a measure of the variability among all possible samples. Although the degree of variability is not known exactly, it can be estimated from the survey results.

The sampling error is usually measured by the *standard error* for a particular statistic (for example, a mean or percentage), which is the square root of the variance. The standard error can be used to calculate confidence intervals within which the true value for the population can reasonably be assumed to fall. For example, for any given statistic calculated from a sample survey, the value of that statistic will fall within a range, calculated as the value of the statistic plus or minus two times the standard error of that statistic, in 95 percent of all possible samples of identical size and design.

If the sample of women had been selected as a simple random sample, it would have been possible, for many statistics, to use straightforward formulas for calculating sampling errors. However, the NFHS-2 sample is the result of a multi-stage stratified sample design, and it is, therefore, necessary to use more complex formulas. The computer software used to calculate sampling errors for NFHS-2 is ISSA (the Integrated System for Survey Analysis). The linear Taylor series approximation method for variance estimation is used for estimates of means, proportions, and ratios. The JACKKNIFE repeated replication method is used with ISSA for variance estimation for more complex statistics such as fertility and mortality rates.

The ISSA package treats any percentage or average as a ratio estimate, $r = y/x$, where y represents the sample value for variable y, and x represents the number of cases in the group or subgroup under consideration. The variance of r is computed using the formula given below, with the standard error being the square root of the variance:

$$var(r) = \frac{1 - f}{x^2} \sum_{h=1}^{H} \left[\frac{m_h}{m_h - 1} \left(\sum_{i=1}^{m_h} z_{hi}^2 - \frac{z_h^2}{m_h} \right) \right]$$

in which

$$z_{hi} = y_{hi} - rx_{hi}$$

$$z_h = y_h - rx_h$$

where

h represents the stratum that varies from 1 to H,

m_h is the total number of PSUs selected in the h^{th} stratum,

y_{hi} is the sum of the values of variable y in PSU i in the h^{th} stratum,

x_{hi} is the sum of the number of cases in PSU i in the h^{th} stratum,

f is the overall sampling fraction, which is so small that the program ignores it.

In addition to the standard error, ISSA computes the relative standard error, confidence limits for the estimates, and the design effect (DEFT) for each estimate. The design effect is defined as the ratio of the standard error using the given sample design to the standard error that would result if a simple random sample had been used. A DEFT value of 1.0 indicates that the sample design is as efficient as a simple random sample, while a value greater than 1.0 indicates the increase in the sampling error due to the use of a more complex and less statistically efficient design.

Sampling errors for NFHS-2 are calculated for selected variables considered to be of primary interest. The results in this appendix are presented for the state as a whole and for urban and rural areas separately, except for the variable on salt iodization for which the results are shown separately for small cities, towns, and rural areas. For each variable, the type of statistic (mean, proportion, ratio, or rate) and the base population are given in Table A.1. Table A.2 presents the value of the statistic (R), its standard error (SE), the relative standard error (SE/R), and the 95 percent confidence limits (R±2SE) for each variable. In addition, for all variables except the fertility and mortality rates, the table shows the unweighted number of cases (N), the weighted number of cases (WN), the standard error assuming a simple random sample (SER), and the design effect (DEFT).

Table A.1 List of selected variables for sampling errors, Haryana, 1998–99

Variable	Estimate	Base population
Sex ratio	Ratio	*De facto* household population
Illiterate	Proportion	*De facto* household population age 6 and above
Have tuberculosis	Rate	1,000 *de jure* household population
Salt iodized at 15 ppm or more	Proportion	Households
Illiterate	Proportion	Ever-married women age 15–49
High school complete and above	Proportion	Ever-married women age 15–49
Currently married	Proportion	Ever-married women age 15–49
Number of children ever born	Mean	Currently married women age 15–49
Number of living children	Mean	Currently married women age 15–49
Have ever used any method	Proportion	Currently married women age 15–49
Currently using any method	Proportion	Currently married women age 15–49
Currently using any modern method	Proportion	Currently married women age 15–49
Currently using pills	Proportion	Currently married women age 15–49
Currently using IUD	Proportion	Currently married women age 15–49
Currently using condoms	Proportion	Currently married women age 15–49
Currently using female sterilization	Proportion	Currently married women age 15–49
Currently using male sterilization	Proportion	Currently married women age 15–49
Currently using rhythm/safe period	Proportion	Currently married women age 15–49
Using public source for modern method	Proportion	Current users of modern methods
Do not want any more children	Proportion	Currently married women age 15–49
Want to delay birth at least 2 years	Proportion	Currently married women age 15–49
Ideal number of children	Mean	Ever-married women age 15–49
Ideal number of sons	Mean	Ever-married women age 15–49
Ideal number of daughters	Mean	Ever-married women age 15–49
Visited by a health/family planning worker	Proportion	Ever-married women age 15–49
Received no antenatal check-up	Proportion	Births in the past 3 years
Received iron and folic acid tablets or syrup	Proportion	Births in the past 3 years
Received medical assistance during delivery	Proportion	Births in the past 3 years
Received postpartum check-up	Proportion	Noninstitutional births in the past 3 years
Had diarrhoea in the past 2 weeks	Proportion	Children under 3 years
Treated with ORS packets	Proportion	Children under 3 with diarrhoea in past 2 weeks
Taken to a health facility/provider for diarrhoea	Proportion	Children under 3 with diarrhoea in past 2 weeks
Showing a vaccination card	Proportion	Children age 12–23 months
Received BCG vaccination	Proportion	Children age 12–23 months
Received DPT vaccination (3 doses)	Proportion	Children age 12–23 months
Received polio vaccination (3 doses)	Proportion	Children age 12–23 months
Received measles vaccination	Proportion	Children age 12–23 months
Fully vaccinated	Proportion	Children age 12–23 months
Received vitamin A	Proportion	Children age 12–35 months
Had reproductive health problem	Proportion	Currently married women age 15–49
Not involved in any decisionmaking	Proportion	Ever-married women age 15–49
Ever beaten or physically mistreated since age15	Proportion	Ever-married women age 15–49
Not worked in past 12 months	Proportion	Ever-married women age 15–49
Anaemic women	Proportion	Ever-married women age 15–49
Anaemic children	Proportion	Children age 6–35 months
Fertility rates	Rate	All women, population
Mortality rates	Rate	Births, population

Table A.2 Sampling errors, Haryana, 1998–99

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Sex ratio (*De facto* household population)									
Urban	868	19.969	2410	2423	18.742	1.065	0.023	828	908
Rural	874	12.584	6047	6030	12.728	0.989	0.014	849	899
Total	872	10.630	8457	8453	10.550	1.008	0.012	851	893
Illiterate (*De facto* household population age 6 and above)									
Urban	0.172	0.022	3965	3989	0.008	2.706	0.126	0.129	0.216
Rural	0.370	0.013	9680	9652	0.006	2.257	0.036	0.343	0.396
Total	0.312	0.014	13645	13641	0.005	2.850	0.046	0.283	0.341
Have tuberculosis (1,000 *de jure* household population)									
Urban	3.050	0.878	4511	4537	0.817	1.075	0.288	1.294	4.807
Rural	3.793	0.688	11281	11249	0.573	1.201	0.181	2.417	5.168
Total	3.580	0.550	15792	15786	0.471	1.168	0.154	2.479	4.680
Salt iodized at 15 ppm or more (Households)									
Small city	0.884	0.029	593	596	0.013	2.228	0.033	0.825	0.942
Town	0.933	0.022	298	301	0.015	1.488	0.023	0.889	0.976
Rural	0.623	0.022	1950	1944	0.011	1.986	0.035	0.579	0.667
Total	0.710	0.020	2841	2841	0.009	2.394	0.029	0.670	0.751
Illiterate (Ever-married women age 15–49)									
Urban	0.295	0.043	826	837	0.016	2.728	0.147	0.208	0.382
Rural	0.656	0.021	2082	2071	0.010	1.992	0.032	0.614	0.697
Total	0.552	0.025	2908	2908	0.009	2.722	0.045	0.502	0.602
High school complete and above (Ever-married women age 15–49)									
Urban	0.462	0.054	826	837	0.017	3.085	0.116	0.355	0.569
Rural	0.095	0.011	2082	2071	0.006	1.639	0.111	0.074	0.116
Total	0.200	0.024	2908	2908	0.007	3.216	0.119	0.153	0.248
Currently married (Ever-married women age 15–49)									
Urban	0.973	0.006	826	837	0.006	1.102	0.006	0.960	0.985
Rural	0.959	0.005	2082	2071	0.004	1.053	0.005	0.949	0.968
Total	0.963	0.004	2908	2908	0.004	1.078	0.004	0.955	0.970
Number of children ever born (Currently married women age 15–49)									
Urban	2.674	0.077	803	814	0.056	1.375	0.029	2.521	2.828
Rural	3.125	0.063	1996	1985	0.044	1.425	0.020	2.999	3.252
Total	2.994	0.054	2799	2799	0.036	1.523	0.018	2.885	3.103
Number of living children (Currently married women age 15–49)									
Urban	2.476	0.062	803	814	0.051	1.219	0.025	2.353	2.599
Rural	2.798	0.056	1996	1985	0.039	1.454	0.020	2.686	2.911
Total	2.705	0.046	2799	2799	0.031	1.477	0.017	2.612	2.797
Have ever used any method (Currently married women age 15–49)									
Urban	0.783	0.023	803	814	0.015	1.593	0.030	0.736	0.829
Rural	0.669	0.017	1996	1985	0.011	1.622	0.026	0.635	0.703
Total	0.702	0.015	2799	2799	0.009	1.704	0.021	0.673	0.732

Table A.2 Sampling errors, Haryana, 1998–99 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases: Unweighted (N)	Number of cases: Weighted (WN)	Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits: R-2SE	Confidence limits: R+2SE
Currently using any method (Currently married women age 15–49)									
Urban	0.672	0.025	803	814	0.017	1.524	0.038	0.621	0.722
Rural	0.604	0.017	1996	1985	0.011	1.585	0.029	0.570	0.639
Total	0.624	0.015	2799	2799	0.009	1.588	0.023	0.595	0.653
Currently using any modern method (Currently married women age 15–49)									
Urban	0.534	0.025	803	814	0.018	1.403	0.046	0.485	0.584
Rural	0.531	0.018	1996	1985	0.011	1.615	0.034	0.495	0.567
Total	0.532	0.015	2799	2799	0.009	1.549	0.027	0.503	0.561
Currently using pills (Currently married women age 15–49)									
Urban	0.023	0.005	803	814	0.005	1.042	0.241	0.012	0.034
Rural	0.020	0.003	1996	1985	0.003	1.053	0.165	0.013	0.027
Total	0.021	0.003	2799	2799	0.003	1.045	0.135	0.015	0.026
Currently using IUD (Currently married women age 15–49)									
Urban	0.059	0.009	803	814	0.008	1.119	0.158	0.040	0.078
Rural	0.026	0.004	1996	1985	0.004	1.075	0.147	0.018	0.034
Total	0.036	0.004	2799	2799	0.004	1.151	0.113	0.028	0.044
Currently using condoms (Currently married women age 15–49)									
Urban	0.148	0.025	803	814	0.013	2.024	0.171	0.098	0.199
Rural	0.034	0.005	1996	1985	0.004	1.186	0.141	0.025	0.044
Total	0.068	0.010	2799	2799	0.005	2.034	0.143	0.048	0.087
Currently using female sterilization (Currently married women age 15–49)									
Urban	0.280	0.020	803	814	0.016	1.280	0.073	0.239	0.320
Rural	0.430	0.018	1996	1985	0.011	1.596	0.041	0.395	0.466
Total	0.387	0.016	2799	2799	0.009	1.685	0.040	0.356	0.418
Currently using male sterilization (Currently married women age 15–49)									
Urban	0.025	0.005	803	814	0.005	1.003	0.223	0.014	0.036
Rural	0.020	0.003	1996	1985	0.003	1.033	0.162	0.014	0.027
Total	0.021	0.003	2799	2799	0.003	1.021	0.131	0.016	0.027
Currently using rhythm/safe period (Currently married women age 15–49)									
Urban	0.059	0.010	803	814	0.008	1.191	0.167	0.040	0.079
Rural	0.039	0.005	1996	1985	0.004	1.206	0.135	0.028	0.049
Total	0.045	0.005	2799	2799	0.004	1.215	0.106	0.035	0.054
Using public source for modern method (Current users of modern methods)									
Urban	0.592	0.037	430	435	0.024	1.578	0.063	0.517	0.667
Rural	0.879	0.011	1062	1055	0.010	1.124	0.013	0.857	0.902
Total	0.795	0.020	1492	1490	0.010	1.904	0.025	0.756	0.835
Do not want any more children (Currently married women age 15–49)									
Urban	0.473	0.029	803	814	0.018	1.657	0.062	0.415	0.532
Rural	0.268	0.012	1996	1985	0.010	1.176	0.043	0.245	0.292
Total	0.328	0.015	2799	2799	0.009	1.689	0.046	0.298	0.358
Want to delay birth at least two years (Currently married women age 15–49)									
Urban	0.082	0.014	803	814	0.010	1.452	0.171	0.054	0.110
Rural	0.101	0.007	1996	1985	0.007	1.064	0.071	0.087	0.116
Total	0.096	0.006	2799	2799	0.006	1.166	0.068	0.083	0.109

Table A.2 Sampling errors, Haryana, 1998–99 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Ideal number of children (Ever-married women age 15–49)									
Urban	2.306	0.052	809	820	0.025	2.038	0.023	2.203	2.410
Rural	2.635	0.042	2029	2018	0.020	2.040	0.016	2.552	2.718
Total	2.540	0.036	2838	2838	0.017	2.191	0.014	2.468	2.613
Ideal number of sons (Ever-married women age 15–49)									
Urban	1.126	0.039	809	820	0.024	1.621	0.034	1.049	1.204
Rural	1.482	0.032	2028	2017	0.017	1.805	0.021	1.419	1.545
Total	1.379	0.030	2837	2837	0.015	2.058	0.022	1.319	1.439
Ideal number of daughters (Ever-married women age 15–49)									
Urban	0.833	0.028	809	820	0.018	1.579	0.033	0.777	0.888
Rural	0.936	0.017	2028	2017	0.012	1.384	0.018	0.902	0.970
Total	0.906	0.015	2837	2837	0.010	1.519	0.017	0.876	0.937
Visited by a health/family planning worker (Ever-married women age 15–49)									
Urban	0.022	0.008	826	837	0.005	1.551	0.362	0.006	0.038
Rural	0.016	0.003	2082	2071	0.003	1.196	0.207	0.009	0.022
Total	0.018	0.003	2908	2908	0.002	1.335	0.185	0.011	0.024
Received no antenatal check-up (Births in past 3 years)									
Urban	0.216	0.043	242	246	0.029	1.484	0.198	0.131	0.302
Rural	0.480	0.030	818	814	0.019	1.544	0.062	0.421	0.540
Total	0.419	0.027	1060	1060	0.017	1.625	0.064	0.365	0.473
Received iron and folic acid tablets or syrup (Births in past 3 years)									
Urban	0.743	0.038	242	246	0.028	1.365	0.052	0.666	0.820
Rural	0.648	0.028	818	814	0.017	1.656	0.043	0.592	0.703
Total	0.670	0.023	1060	1060	0.014	1.616	0.035	0.623	0.716
Received medical assistance during delivery (Births in past 3 years)									
Urban	0.661	0.048	242	246	0.032	1.479	0.073	0.565	0.757
Rural	0.347	0.025	818	814	0.018	1.335	0.071	0.298	0.397
Total	0.420	0.025	1060	1060	0.017	1.516	0.060	0.370	0.470
Received postpartum check-up (Noninstitutional births in past 3 years)									
Urban	0.218	0.049	128	130	0.037	1.325	0.223	0.121	0.315
Rural	0.146	0.018	696	693	0.013	1.324	0.121	0.111	0.181
Total	0.157	0.017	824	823	0.013	1.335	0.108	0.123	0.191
Had diarrhoea in the past 2 weeks (Children under 3 years)									
Urban	0.127	0.024	230	234	0.022	1.084	0.188	0.079	0.174
Rural	0.142	0.013	750	747	0.013	1.052	0.094	0.116	0.169
Total	0.139	0.012	980	980	0.011	1.056	0.084	0.115	0.162
Treated with ORS packets (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.374	0.096	29	30	0.090	1.072	0.257	0.182	0.566
Rural	0.225	0.043	107	106	0.040	1.069	0.193	0.138	0.312
Total	0.257	0.040	136	136	0.038	1.075	0.157	0.177	0.338
Taken to a health facility/provider for diarrhoea (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.896	0.059	29	30	0.056	1.049	0.066	0.778	1.000
Rural	0.934	0.023	107	106	0.024	0.953	0.025	0.888	0.980
Total	0.926	0.022	136	136	0.023	0.964	0.023	0.882	0.969

Table A.2 Sampling errors, Haryana, 1998–99 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Showing a vaccination card (Children age 12–23 months)									
Urban	0.433	0.052	81	83	0.056	0.938	0.121	0.328	0.537
Rural	0.182	0.030	253	252	0.024	1.236	0.165	0.122	0.242
Total	0.244	0.028	334	335	0.024	1.197	0.116	0.187	0.301
Received BCG vaccination (Children age 12–23 months)									
Urban	0.890	0.035	81	83	0.035	1.008	0.039	0.820	0.960
Rural	0.861	0.028	253	252	0.022	1.298	0.033	0.805	0.918
Total	0.868	0.023	334	335	0.018	1.232	0.026	0.823	0.914
Received DPT vaccination (3 doses) (Children age 12–23 months)									
Urban	0.840	0.043	81	83	0.041	1.059	0.051	0.753	0.926
Rural	0.669	0.037	253	252	0.030	1.247	0.055	0.595	0.742
Total	0.711	0.030	334	335	0.025	1.218	0.043	0.650	0.771
Received polio vaccination (3 doses) (Children age 12–23 months)									
Urban	0.852	0.040	81	83	0.039	1.014	0.047	0.773	0.932
Rural	0.707	0.035	253	252	0.029	1.210	0.049	0.638	0.776
Total	0.743	0.028	334	335	0.024	1.180	0.038	0.687	0.799
Received measles vaccination (Children age 12–23 months)									
Urban	0.852	0.044	81	83	0.039	1.118	0.052	0.764	0.940
Rural	0.680	0.034	253	252	0.029	1.165	0.050	0.612	0.748
Total	0.722	0.028	334	335	0.025	1.156	0.039	0.666	0.779
Fully vaccinated (Children age 12–23 months)									
Urban	0.765	0.051	81	83	0.047	1.072	0.066	0.663	0.866
Rural	0.582	0.038	253	252	0.031	1.227	0.065	0.506	0.658
Total	0.627	0.031	334	335	0.026	1.181	0.050	0.565	0.690
Received vitamin A (Children age 12–35 months)									
Urban	0.544	0.043	159	162	0.041	1.036	0.079	0.458	0.629
Rural	0.422	0.030	502	500	0.022	1.327	0.070	0.363	0.481
Total	0.452	0.025	661	661	0.020	1.264	0.055	0.402	0.502
Had reproductive health problem (Currently married women age 15–49)									
Urban	0.341	0.018	803	814	0.017	1.084	0.053	0.305	0.377
Rural	0.399	0.014	1996	1985	0.011	1.252	0.034	0.372	0.427
Total	0.382	0.011	2799	2799	0.009	1.223	0.029	0.360	0.405
Not involved in any decisionmaking (Ever-married women age 15–49)									
Urban	0.037	0.008	826	837	0.007	1.268	0.225	0.020	0.054
Rural	0.033	0.006	2082	2071	0.004	1.494	0.176	0.022	0.045
Total	0.034	0.005	2908	2908	0.003	1.421	0.140	0.025	0.044
Ever beaten or physically mistreated since age 15 (Ever-married women age 15–49)									
Urban	0.117	0.013	826	837	0.011	1.202	0.115	0.091	0.144
Rural	0.138	0.011	2082	2071	0.008	1.466	0.080	0.116	0.160
Total	0.132	0.009	2908	2908	0.006	1.399	0.067	0.114	0.149
Not worked in past 12 months (Ever-married women age 15–49)									
Urban	0.849	0.015	826	837	0.012	1.235	0.018	0.818	0.880
Rural	0.884	0.010	2082	2071	0.007	1.459	0.012	0.864	0.905
Total	0.874	0.009	2908	2908	0.006	1.413	0.010	0.857	0.892

Table A.2 Sampling errors, Haryana, 1998–99 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Anaemic women (Ever-married women age 15–49)									
Urban	0.458	0.025	761	787	0.018	1.393	0.055	0.408	0.508
Rural	0.475	0.014	1974	1948	0.011	1.246	0.030	0.447	0.503
Total	0.470	0.012	2735	2735	0.010	1.290	0.026	0.445	0.495
Anaemic children (Children age 6–35 months)									
Urban	0.866	0.030	175	182	0.026	1.170	0.035	0.806	0.927
Rural	0.830	0.016	562	555	0.016	1.040	0.020	0.797	0.863
Total	0.839	0.014	737	737	0.014	1.065	0.017	0.810	0.868

Table A.2 Sampling errors, Haryana, 1998–99 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Relative standard error (SE/R)	Confidence limits	
				R-2SE	R+2SE
Total fertility rate (Women age 15–49)					
Urban	2.243	0.141	0.063	1.962	2.525
Rural	3.132	0.126	0.040	2.880	3.384
Total	2.879	0.097	0.034	2.686	3.072
Age-specific fertility rate (Women age 15–19)					
Urban	0.031	0.007	0.225	0.017	0.044
Rural	0.116	0.012	0.103	0.092	0.140
Total	0.092	0.009	0.096	0.074	0.110
Age-specific fertility rate (Women age 20–24)					
Urban	0.186	0.014	0.073	0.159	0.214
Rural	0.260	0.011	0.042	0.239	0.282
Total	0.240	0.009	0.036	0.222	0.257
Age-specific fertility rate (Women age 25–29)					
Urban	0.151	0.014	0.091	0.124	0.179
Rural	0.150	0.010	0.067	0.130	0.170
Total	0.150	0.008	0.054	0.134	0.167
Age-specific fertility rate (Women age 30–34)					
Urban	0.063	0.014	0.218	0.036	0.090
Rural	0.062	0.009	0.147	0.043	0.080
Total	0.062	0.007	0.120	0.047	0.077
Age-specific fertility rate (Women age 35–39)					
Urban	0.011	0.007	0.050	0.003	0.026
Rural	0.018	0.007	0.410	0.003	0.032
Total	0.015	0.005	0.344	0.005	0.026
Age-specific fertility rate (Women age 40–44)					
Urban	0.000	0.000	NC	0.000	0.000
Rural	0.013	0.005	0.417	0.002	0.024
Total	0.009	0.004	0.421	0.001	0.017
Age-specific fertility rate (Women age 45–49)					
Urban	0.006	0.006	1.027	0.006	0.019
Rural	0.008	0.006	0.752	0.004	0.020
Total	0.007	0.004	0.614	0.002	0.016

Table A.2 Sampling errors, Haryana, 1999 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Relative standard error (SE/R)	Confidence limits	
				R-2SE	R+2SE
Neonatal mortality (5-year period preceding survey)					
Urban	32.559	9.145	0.281	14.269	50.849
Rural	35.588	5.082	0.143	25.423	45.752
Total	34.870	4.397	0.126	26.076	43.664
Infant mortality $_1q_0$ (5-year period preceding survey)					
Urban	41.638	10.022	0.241	21.594	61.683
Rural	61.508	6.802	0.111	47.903	75.113
Total	56.767	5.719	0.101	45.329	68.205
Child mortality $_4q_1$ (5-year period preceding survey)					
Urban	19.606	6.575	0.335	6.457	32.755
Rural	21.738	4.227	0.194	13.284	30.192
Total	21.197	3.530	0.167	14.136	28.257
Under-five mortality $_5q_0$ (5-year period preceding survey)					
Urban	60.428	11.518	0.191	37.392	83.464
Rural	81.909	7.155	0.087	67.599	96.219
Total	76.761	6.111	0.080	64.539	88.982
Crude death rate (Based on Household Questionnaire)					
Urban	7.444	1.237	0.166	4.971	9.918
Rural	8.460	0.696	0.082	7.068	9.851
Total	8.168	0.628	0.077	6.912	9.423
Crude birth rate (Based on women's birth history)					
Urban	18.135	1.117	0.062	15.902	20.368
Rural	24.995	0.995	0.040	23.005	26.984
Total	23.117	0.778	0.034	21.562	24.672

NC: Not calculated because denominator is 0.000
SRS: Simple random sample

APPENDIX B

DATA QUALITY TABLES

The purpose of this appendix is to provide the data user with an overview of the general quality of the NFHS-2 data. Whereas Appendix A is concerned with sampling errors and their effects on the survey results, the tables in this appendix refer to possible *nonsampling* errors: for example, rounding or heaping on certain ages or dates; omission of events occurring further in the past; deliberate distortion of information by some interviewers in an attempt to lighten their workload; noncooperation of the respondent in providing information; or refusal to have children measured for height and weight or tested for anaemia. A description of the likely magnitude of such nonsampling errors is provided in this appendix.

The distribution of the *de facto* household population by single years of age and sex is presented in Table B.1. In many (but not all) cases, the respondent was the head of the household. It is well documented that ages are poorly reported in most parts of India. Ages are of little relevance to much of the rural population in particular, and no amount of probing will ensure that ages are properly recorded. In interviewer training for NFHS-2, a great deal of emphasis was placed on obtaining as accurate information as possible on ages and dates of events. Nevertheless, it is clear that age reporting in NFHS-2 shares the same problems inherent in all Indian censuses and surveys. Heaping on ages ending in 0, 2, 5, and 8 is considerable and is particularly severe in the older age groups. However, the NFHS-2 age data are evidently of considerably better quality than age data from other sources. This can be seen, for example, by comparing the degree of age heaping in NFHS-2 with that in the 1991 Census. Another measure of the quality of the NFHS-2 age data is the percentage of persons whose ages were recorded as not known or missing. In Haryana, information on age was missing for only 1 person out of 15,825 persons who stayed in the sample households the night before the interview.

Table B.2 examines the possibility that some eligible women (that is, ever-married women age 15–49) were not properly identified in NFHS-2. In some surveys, interviewers may try to reduce their workload by pushing women out of the eligible age range or recording ever-married women as never married so that they will not have to be interviewed. If such practices were being followed to a noticeable extent, Table B.2 would normally show (1) a shortage of ever-married women in the 45–49 age group and an excess in the 50–54 age group or (2) an unusually low proportion of ever-married women by age. Neither of these patterns is evident in the NFHS-2 data. It can, therefore, be concluded that there was no concerted effort to misidentify eligible women in NFHS-2 in Haryana.

One traditional measure of the quality of data is the extent to which information is missing on key variables. Although completeness of responses does not necessarily indicate that the results are accurate, the existence of missing information for a large number of cases would suggest that data collection was not carried out with sufficient care. In NFHS-2 in Haryana, the extent of missing information is very low for month and year of birth, age at death, age at first marriage, woman's education, and prevalence of diarrhoea in the two weeks preceding the survey (Table B.3). Data on height and weight of children are available for 96 percent of cases, and data on haemoglobin levels for women and children are available for 92–94 percent of cases. The high response rates for anthropometric and haemoglobin measurements are impressive.

Table B.1 Household age distribution

Single-year age distribution of *de facto* household population by sex (weighted), Haryana, 1998–99

	Male		Female			Male		Female	
Age	Number	Percent	Number	Percent	Age	Number	Percent	Number	Percent
< 1	199	2.4	161	2.2	38	103	1.2	101	1.4
1	177	2.1	145	2.0	39	22	0.3	30	0.4
2	184	2.2	155	2.1	40	297	3.5	205	2.8
3	204	2.4	153	2.1	41	15	0.2	22	0.3
4	187	2.2	197	2.7	42	70	0.8	66	0.9
5	230	2.7	193	2.6	43	18	0.2	28	0.4
6	232	2.7	170	2.3	44	14	0.2	23	0.3
7	194	2.3	182	2.5	45	225	2.7	165	2.2
8	269	3.2	186	2.5	46	28	0.3	40	0.5
9	211	2.5	155	2.1	47	20	0.2	30	0.4
10	260	3.1	236	3.2	48	58	0.7	62	0.8
11	168	2.0	149	2.0	49	8	0.1	34	0.5
12	252	3.0	225	3.0	50	160	1.9	31	0.4
13	180	2.1	174	2.4	51	14	0.2	16	0.2
14	192	2.3	184	2.5	52	42	0.5	69	0.9
15	200	2.4	184	2.5	53	13	0.2	45	0.6
16	223	2.6	173	2.4	54	14	0.2	26	0.4
17	151	1.8	113	1.5	55	113	1.3	119	1.6
18	253	3.0	191	2.6	56	31	0.4	23	0.3
19	128	1.5	85	1.2	57	15	0.2	11	0.1
20	221	2.6	201	2.7	58	45	0.5	49	0.7
21	108	1.3	97	1.3	59	7	0.1	10	0.1
22	225	2.7	171	2.3	60	145	1.7	153	2.1
23	112	1.3	112	1.5	61	10	0.1	7	0.1
24	120	1.4	78	1.1	62	34	0.4	39	0.5
25	219	2.6	238	3.2	63	18	0.2	6	0.1
26	117	1.4	117	1.6	64	9	0.1	3	0.0
27	90	1.1	72	1.0	65	136	1.6	119	1.6
28	150	1.8	175	2.4	66	9	0.1	13	0.2
29	42	0.5	39	0.5	67	7	0.1	4	0.1
30	282	3.3	274	3.7	68	23	0.3	24	0.3
31	27	0.3	27	0.4	69	3	0.0	9	0.1
32	128	1.5	116	1.6	70+	314	3.7	207	2.8
33	33	0.4	55	0.7	Don't				
34	46	0.5	49	0.7	know/				
35	312	3.7	255	3.5	missing	0	0.0	1	0.0
36	62	0.7	70	1.0					
37	23	0.3	23	0.3	Total	8,453	100.0	7,372	100.0

Note: The *de facto* population includes both usual residents and visitors who stayed in the household the night before the interview.

Some children could not be measured because they were not at home or they were ill at the time of the survey. In some other cases when the child was at home and not ill, either the child refused to be measured or the mother refused to allow the child to be measured. Before undertaking haemoglobin measurements, a separate 'informed consent' statement was read to the respondent explaining that participation in the haemoglobin testing was completely voluntary. At this point, some women declined to take part in the anaemia testing and/or to have their children participate.

Another measure of data quality is the completeness and accuracy of information on births. Table B.4 examines the distribution of births by calendar year to identify any unusual patterns that may indicate that births have been omitted or that the ages of children have been displaced. All living children listed in the birth history had complete birth dates recorded, as did 99.8 percent of children who had died. The completeness of data on birth dates for both surviving and nonsurving children is exceptionally good. The annual data on the number of births can be examined to see if there is an abnormally large decline in the number of births after

Table B.2 Age distribution of eligible and interviewed women

Age distribution of the *de facto* household population of women age 10–54 and of interviewed women age 15–49, and percentage of eligible women who were interviewed (weighted), Haryana, 1998–99

Age	All women	Ever-married women	Interviewed women		Percent interviewed
			Number	Percent	
10–14	968	0	NA	NA	NA
15–19	746	167	162	5.6	97.0
20–24	659	516	509	17.5	98.7
25–29	640	623	606	20.9	97.3
30–34	521	521	512	17.6	98.1
35–39	480	478	463	15.9	96.8
40–44	344	343	338	11.6	98.5
45–49	331	329	317	10.9	96.3
50–54	187	186	NA	NA	NA
15–49	3,722	2,978	2,907	100.0	97.6

Note: The *de facto* population includes both usual residents and visitors who stayed in the household the night before the interview. For all columns, the age distribution is taken from ages reported in the Household Questionnaire. The total number of interviewed women in this table differs from the total number in earlier tables because this table uses household weights rather than women's weights for the calculations.
NA: Not applicable

Table B.3 Completeness of reporting

Percentage of observations with missing information for selected demographic and health indicators (weighted), Haryana, 1998–99

Indicator	Reference group	Percentage missing information	Number of cases
Birth date	Births in past 15 years		
Month only		0.05	5,870
Month and year		0.00	5,870
Age at death	Deaths to births in past 15 years	0.20	509
Age at first marriage	Ever-married women age 15–49	0.07	2,908
Woman's education	Ever-married women age 15–49	0.04	2,908
Anthropometry	Living children age 0–35 months		
Height		4.46	995
Weight		4.16	995
Height or weight		4.46	995
Woman's haemoglobin level	Ever-married women age 15–49	5.98	2,908
Child's haemoglobin level	Living children age 6–35 months	8.15	802
Diarrhoea in past 2 weeks	Living children age 1–35 months	0.10	980

Table B.4 Births by calendar year

Number of births, percent with complete birth date, sex ratio at birth, and calendar year ratio for children still alive at the time of the survey (L), children who died by the time of the survey (D), and total children (T), by calendar year (weighted), Haryana, 1998–99

	Number of births			Percent with complete birth date[1]			Sex ratio at birth[2]			Calendar year ratio[3]		
Calendar year	L	D	T	L	D	T	L	D	T	L	D	T
1999	24	1	25	100.0	100.0	100.0	848	NC	924	NA	NA	NA
1998	338	22	360	100.0	100.0	100.0	867	824	865	NC	NC	NC
1997	322	20	342	100.0	100.0	100.0	796	1,004	807	NC	NC	NC
1996	330	26	356	100.0	100.0	100.0	786	732	782	99.5	144.8	101.8
1995	340	16	356	100.0	100.0	100.0	809	1,007	817	96.3	51.5	92.7
1994	377	36	413	100.0	100.0	100.0	966	1,412	998	104.3	134.0	106.4
1993	382	38	420	100.0	100.0	100.0	905	1,901	966	98.9	105.1	99.4
1992	396	36	432	100.0	100.0	100.0	785	1,385	823	107.3	101.3	106.7
1991	356	33	389	100.0	100.0	100.0	892	2,307	963	89.1	93.5	89.4
1990	404	35	439	100.0	100.0	100.0	710	835	719	114.4	106.9	113.8
1989	350	32	382	100.0	100.0	100.0	738	1,295	774	83.2	82.2	83.1
1988	438	43	481	99.8	100.0	99.8	969	1,271	993	133.7	124.9	132.9
1993–97	1,751	136	1,886	100.0	100.0	100.0	854	1,229	877	NA	NA	NA
1988–92	1,945	179	2,124	99.9	100.0	100.0	816	1,325	851	NA	NA	NA
1983–87	1,622	203	1,825	100.0	99.0	99.9	909	827	900	NA	NA	NA
1978–82	1,271	163	1,434	100.0	100.0	100.0	725	1,034	755	NA	NA	NA
1977 or earlier	959	159	1,118	100.0	100.0	100.0	906	959	914	NA	NA	NA
All	7,909	863	8,773	100.0	99.8	100.0	840	1,043	858	NA	NA	NA

NA: Not applicable
NC: Not calculated because full-year data were not collected for 1998 and 1999 (the survey began during 1998) or the denominator is zero (for children born in 1999 who died)
[1]Both year and month of birth given
[2]$(B_f/B_m)x1000$, where B_f and B_m are the numbers of female and male births, respectively
[3]$[2B_x/(B_{x-1}+B_{x+1})]x100$, where B_x is the number of births in calendar year x

January 1995, the cutoff point for the health questions and measurements made on young children in the survey. It is typical for the annual number of births to fluctuate somewhat, so small annual fluctuations are to be expected. However, a noticeable drop in the annual number of births between 1992–94 and 1995–98 (particularly for nonsurviving children) suggests that there has been some omission of recent births or displacement of birth dates that could result in an underestimate of both fertility and infant mortality rates for recent years.

Many surveys that include both demographic information and health information for children below a specified age have been subject to a substantial amount of age displacement. In particular, there is often a tendency for interviewers to 'age' children out of the eligible period for asking health questions. This problem was well known before NFHS-2 began; therefore, interviewer training stressed this issue to try to reduce the extent of biases due to age displacement. Apparently, the training was not entirely successful in avoiding this type of problem, however.

Table B.5 presents information on the reporting of age at death in days. Results from the table suggest that early infant deaths have not been seriously underreported in Haryana because the ratios of deaths under seven days to all neonatal deaths are consistently high (a ratio of less than 25 percent is often used as a guideline to indicate underreporting of early neonatal deaths). The ratios are 67 for 0–4 years, 69 for 5–9 years, and 72 for 10–14 years preceding the survey. Although there was no severe underreporting of early neonatal deaths in NFHS-2, there was some misreporting of age at death due to a preference for reporting the age at death at 3, 6, 8, and 15 days (Table B.5).

Table B.6 shows the percentage of infant deaths that occurred during the neonatal period. These percentages are also quite high, suggesting that there is no major omission of early deaths. One problem that is inherent in most retrospective surveys is heaping of the age at death on certain digits, e.g., 6, 12, and 18 months. Misreporting of age at death will bias estimates of the age pattern of mortality if the net result of misreporting is the transference of deaths between age segments for which the rates are calculated. For example, an overestimate of child mortality relative to infant mortality may result if children dying during the first year of life are reported as having died at age one or older. Thus, heaping at 12 months can bias the mortality estimates because a certain fraction of these deaths may have actually occurred during infancy (that is, at ages 0–11 months). In this case, heaping would bias the infant mortality rate downward and child mortality upward.

Examination of the distribution of deaths under age two years during the 15 years before the survey by month of death (Table B.6) indicates there is substantial heaping of deaths at 6, 12, and 18 months of age. Digit preference appears not to be serious enough to alter substantially the mortality rates calculated here. Nevertheless, even if one-third of the deaths reported at age 12 months or age one year actually occurred at less than 12 months of age, the infant mortality rate for the five years before the survey would be underestimated by only 2 percent.

This brief check on internal consistency of NFHS-2 childhood mortality data for Haryana suggests that there is no serious underreporting of deaths during the time periods for which mortality rates are estimated. Although there is some heaping of age at death at certain ages, the heaping is minimal and any resulting bias in infant and child mortality rates should be negligible.

Table B.5 Reporting of age at death in days

Distribution of reported deaths under 1 month of age by age at death in days and percentage of neonatal deaths reported to occur at age 0–6 days, for births occurring during five-year periods preceding the survey (weighted), Haryana, 1998–99

	Years preceding survey			
Age at death (days)	0–4	5–9	10–14	0–14
< 1	9	17	13	39
1	15	19	19	53
2	3	1	5	9
3	6	7	3	16
4	5	1	2	8
5	1	1	5	7
6	3	2	9	14
7	1	2	3	6
8	2	2	6	10
9	0	2	2	4
10	0	2	3	5
11	2	2	1	5
12	1	1	1	3
13	1	2	0	3
14	1	1	0	2
15	3	2	2	7
16	1	0	0	1
17	0	0	0	0
18	2	1	0	3
19	1	1	0	2
20	1	1	2	4
21	0	0	0	0
22	1	0	1	2
23	0	0	0	0
24	2	0	1	3
25	2	3	0	5
26	0	0	0	0
27	0	0	0	0
28	0	0	0	0
29	0	0	0	0
30	0	0	0	0
0–30	63	69	78	210
Percent early neonatal[1]	66.7	68.5	71.8	69.2

[1]Deaths during the first 6 days divided by deaths during the first 30 days

Table B.6 Reporting of age at death in months

Distribution of reported deaths under two years of age by age at death in months and percentage of infant deaths reported to occur at age under one month, for births occurring during five-year periods preceding the survey (weighted), Haryana, 1998–99

	Years preceding survey			
Age at death (months)	0–4	5–9	10–14	0–14
< 1	64	69	78	211
1	9	11	12	32
2	3	6	9	18
3	3	7	4	14
4	1	5	6	12
5	2	3	4	9
6	8	7	7	22
7	1	1	2	4
8	3	5	4	12
9	0	4	10	14
10	5	3	2	10
11	2	2	7	11
12	5	16	20	41
13	4	2	1	7
14	1	1	1	3
15	2	2	2	6
16	0	0	0	0
17	0	0	0	0
18	0	2	7	9
19	1	1	0	2
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	0	1	2
1 year	1	2	1	4
0–11 months	101	123	144	368
Percent neonatal[1]	63.5	56.3	53.9	57.3

[1]Deaths during the first month divided by deaths during the first year