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 = the stratum that varies from 1 to H,
m_h = the total number of PSUs selected in the h^{th} stratum,
y_{hi} = the sum of the values of variable y in PSU i in the h^{th} stratum,
x_{hi} = the sum of the number of cases in PSU i in the h^{th} stratum,
f = 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 large cities, 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, Maharashtra, 1999

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 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 age 15	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, Maharashtra, 1999

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
Sex ratio (*De facto* Household population)									
Urban	898	20.381	9471	6673	10.618	1.919	0.023	857	939
Rural	985	14.361	5670	8618	13.980	1.027	0.015	957	1014
Total	947	12.231	15141	15290	8.487	1.441	0.013	923	972
Illiterate (*De facto* Household population age 6 and above)									
Urban	0.173	0.015	15762	11156	0.004	3.749	0.085	0.144	0.203
Rural	0.355	0.016	9729	14789	0.006	2.457	0.044	0.323	0.386
Total	0.277	0.013	25491	25945	0.004	3.339	0.045	0.252	0.302
Have tuberculosis (1,000 *de jure* household population)									
Urban	3.422	0.604	18335	12839	0.498	1.213	0.176	2.215	4.630
Rural	2.359	0.818	11233	17078	0.793	1.031	0.347	0.723	3.996
Total	2.816	0.531	29568	29917	0.448	1.185	0.189	1.754	3.877
Salt iodized at 15 ppm or more (Households)									
Large city	0.679	0.029	2675	1184	0.009	3.183	0.042	0.622	0.737
Small city	0.733	0.057	612	838	0.018	3.193	0.078	0.618	0.847
Town	0.730	0.104	375	510	0.023	4.534	0.143	0.522	0.938
Rural	0.519	0.031	2168	3298	0.011	2.880	0.060	0.457	0.580
Total	0.601	0.023	5830	5830	0.006	3.556	0.038	0.555	0.646
Illiterate (Ever-married women age 15–49)									
Urban	0.290	0.025	3191	2229	0.008	3.170	0.088	0.239	0.341
Rural	0.556	0.024	2200	3162	0.011	2.297	0.044	0.507	0.605
Total	0.446	0.020	5391	5391	0.007	3.010	0.046	0.405	0.487
High school complete and above (Ever-married women age 15–49)									
Urban	0.303	0.027	3191	2229	0.008	3.357	0.090	0.248	0.357
Rural	0.089	0.011	2200	3162	0.006	1.800	0.123	0.067	0.111
Total	0.177	0.015	5391	5391	0.005	2.868	0.084	0.148	0.207
Currently married women (Ever-married women age 15–49)									
Urban	0.917	0.007	3191	2229	0.005	1.409	0.007	0.903	0.931
Rural	0.923	0.007	2200	3162	0.006	1.212	0.007	0.909	0.937
Total	0.921	0.005	5391	5391	0.004	1.345	0.005	0.911	0.931
Number of children ever born (Currently married women age 15–49)									
Urban	2.554	0.062	2951	2044	0.032	1.954	0.024	2.430	2.679
Rural	2.955	0.049	2031	2919	0.039	1.236	0.016	2.858	3.052
Total	2.790	0.040	4982	4963	0.025	1.595	0.014	2.710	2.870
Number of living children (Currently married women age 15–49)									
Urban	2.369	0.053	2951	2044	0.028	1.875	0.022	2.263	2.476
Rural	2.606	0.039	2031	2919	0.033	1.184	0.015	2.528	2.685
Total	2.509	0.032	4982	4963	0.022	1.497	0.013	2.444	2.573
Have ever used any method (Currently married women age 15–49)									
Urban	0.670	0.016	2951	2044	0.009	1.801	0.023	0.639	0.701
Rural	0.659	0.014	2031	2919	0.011	1.321	0.021	0.631	0.687
Total	0.664	0.010	4982	4963	0.007	1.546	0.016	0.643	0.684

Table A.2 Sampling errors, Maharashtra, 1999 (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	0.585	0.017	2951	2044	0.009	1.859	0.029	0.551	0.618
Rural	0.627	0.014	2031	2919	0.011	1.308	0.022	0.599	0.655
Total	0.609	0.011	4982	4963	0.007	1.570	0.018	0.588	0.631
Currently using any modern method (Currently married women age 15–49)									
Urban	0.567	0.017	2951	2044	0.009	1.844	0.030	0.533	0.600
Rural	0.621	0.014	2031	2919	0.011	1.316	0.023	0.593	0.650
Total	0.599	0.011	4982	4963	0.007	1.576	0.018	0.577	0.621
Currently using pills (Currently married women age 15–49)									
Urban	0.025	0.004	2951	2044	0.003	1.383	0.160	0.017	0.033
Rural	0.012	0.003	2031	2919	0.002	1.022	0.202	0.007	0.017
Total	0.017	0.002	4982	4963	0.002	1.214	0.129	0.013	0.022
Currently using IUD (Currently married women age 15–49)									
Urban	0.035	0.006	2951	2044	0.003	1.694	0.164	0.023	0.046
Rural	0.008	0.003	2031	2919	0.002	1.239	0.299	0.003	0.013
Total	0.019	0.003	4982	4963	0.002	1.468	0.148	0.014	0.025
Currently using condoms (Currently married women age 15–49)									
Urban	0.056	0.007	2951	2044	0.004	1.634	0.123	0.042	0.070
Rural	0.029	0.004	2031	2919	0.004	1.174	0.151	0.020	0.038
Total	0.040	0.004	4982	4963	0.003	1.406	0.097	0.032	0.048
Currently using female sterilization (Currently married women age 15–49)									
Urban	0.436	0.017	2951	2044	0.009	1.871	0.039	0.402	0.470
Rural	0.519	0.017	2031	2919	0.011	1.536	0.033	0.485	0.553
Total	0.485	0.012	4982	4963	0.007	1.754	0.026	0.460	0.509
Currently using male sterilization (Currently married women age 15–49)									
Urban	0.015	0.003	2951	2044	0.002	1.504	0.222	0.009	0.022
Rural	0.053	0.009	2031	2919	0.005	1.755	0.165	0.035	0.070
Total	0.037	0.006	4982	4963	0.003	2.059	0.148	0.026	0.048
Currently using rhythm/safe period (Currently married women age 15–49)									
Urban	0.014	0.003	2951	2044	0.002	1.354	0.209	0.008	0.020
Rural	0.002	0.001	2031	2919	0.001	0.977	0.435	0.000	0.005
Total	0.007	0.001	4982	4963	0.001	1.147	0.191	0.004	0.010
Using public source for modern method (Current users of modern methods)									
Urban	0.591	0.032	1630	1159	0.012	2.651	0.055	0.526	0.655
Rural	0.855	0.015	1262	1814	0.010	1.564	0.018	0.824	0.886
Total	0.752	0.018	2892	2972	0.008	2.299	0.025	0.715	0.789
Do not want any more children (Currently married women age 15–49)									
Urban	0.256	0.013	2951	2044	0.008	1.660	0.052	0.229	0.283
Rural	0.158	0.010	2031	2919	0.008	1.213	0.062	0.138	0.177
Total	0.198	0.009	4982	4963	0.006	1.526	0.043	0.181	0.215
Want to delay birth at least two years (Currently married women age 15–49)									
Urban	0.112	0.009	2951	2044	0.006	1.563	0.081	0.094	0.130
Rural	0.082	0.007	2031	2919	0.006	1.193	0.088	0.068	0.097
Total	0.095	0.006	4982	4963	0.004	1.411	0.062	0.083	0.106

Table A.2 Sampling errors, Maharashtra, 1999 (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
Ideal number of children (Ever-married women age 15–49)									
Urban	2.198	0.044	3112	2170	0.013	3.401	0.020	2.110	2.287
Rural	2.388	0.028	2079	2987	0.016	1.735	0.012	2.332	2.443
Total	2.308	0.025	5191	5157	0.010	2.424	0.011	2.259	2.357
Ideal number of sons (Ever-married women age 15–49)									
Urban	0.998	0.036	3112	2170	0.013	2.899	0.036	0.926	1.071
Rural	1.274	0.025	2079	2987	0.015	1.707	0.019	1.224	1.323
Total	1.158	0.023	5191	5157	0.010	2.363	0.020	1.113	1.203
Ideal number of daughters (Ever-married women age 15–49)									
Urban	0.803	0.029	3112	2170	0.010	2.932	0.036	0.745	0.861
Rural	0.909	0.016	2079	2987	0.011	1.446	0.017	0.878	0.940
Total	0.864	0.015	5191	5157	0.007	2.106	0.018	0.834	0.895
Visited by health/family planning worker (Ever-married women age 15–49)									
Urban	0.174	0.019	3191	2229	0.007	2.825	0.109	0.136	0.212
Rural	0.277	0.020	2200	3162	0.010	2.120	0.073	0.237	0.318
Total	0.234	0.015	5391	5391	0.006	2.549	0.063	0.205	0.264
Received no antenatal check-up (Births in the past 3 years)									
Urban	0.052	0.011	993	707	0.007	1.530	0.218	0.029	0.074
Rural	0.125	0.020	768	1103	0.013	1.548	0.160	0.085	0.165
Total	0.096	0.013	1761	1810	0.007	1.775	0.138	0.070	0.123
Received iron and folic acid tablets or syrup (Births in the past 3 years)									
Urban	0.886	0.014	993	707	0.010	1.426	0.016	0.858	0.915
Rural	0.823	0.021	768	1103	0.014	1.516	0.025	0.782	0.865
Total	0.848	0.014	1761	1810	0.009	1.654	0.017	0.820	0.876
Received medical assistance during delivery (Births in the past 3 years)									
Urban	0.841	0.022	993	707	0.012	1.794	0.026	0.797	0.885
Rural	0.436	0.032	768	1103	0.019	1.635	0.073	0.372	0.500
Total	0.594	0.028	1761	1810	0.013	2.228	0.047	0.539	0.650
Received postpartum check-up (Noninstitutional births in the past 3 years)									
Urban	0.393	0.044	167	135	0.038	1.171	0.113	0.304	0.482
Rural	0.280	0.026	503	723	0.020	1.317	0.094	0.227	0.333
Total	0.298	0.024	670	858	0.018	1.349	0.080	0.250	0.345
Had diarrhoea in the past 2 weeks (Children under 3 years)									
Urban	0.291	0.030	949	673	0.015	2.004	0.102	0.232	0.350
Rural	0.229	0.019	717	1030	0.016	1.209	0.083	0.191	0.267
Total	0.254	0.017	1666	1704	0.011	1.582	0.067	0.220	0.287
Treated with ORS packets (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.305	0.046	240	196	0.028	1.629	0.150	0.214	0.397
Rural	0.354	0.047	164	236	0.039	1.211	0.132	0.261	0.448
Total	0.332	0.033	404	432	0.023	1.404	0.099	0.267	0.397
Taken to a health facility/provider for diarrhoea (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.850	0.023	240	196	0.022	1.024	0.027	0.804	0.895
Rural	0.708	0.041	164	236	0.036	1.151	0.059	0.625	0.791
Total	0.772	0.025	404	432	0.021	1.210	0.032	0.722	0.822

Table A.2 Sampling errors, Maharashtra, 1999 (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
Showing a vaccination card (Children age 12–23 months)									
Urban	0.536	0.034	311	229	0.028	1.224	0.063	0.468	0.603
Rural	0.459	0.036	252	362	0.032	1.136	0.078	0.387	0.531
Total	0.489	0.026	563	591	0.021	1.248	0.053	0.437	0.540
Received BCG vaccination (Children age 12–23 months)									
Urban	0.983	0.009	311	229	0.007	1.268	0.009	0.965	1.000
Rural	0.908	0.027	252	362	0.018	1.474	0.030	0.854	0.961
Total	0.937	0.017	563	591	0.010	1.747	0.019	0.902	0.972
Received DPT vaccination (3 doses) (Children age 12–23 months)									
Urban	0.922	0.019	311	229	0.015	1.260	0.020	0.885	0.959
Rural	0.877	0.034	252	362	0.021	1.647	0.039	0.809	0.945
Total	0.894	0.022	563	591	0.013	1.756	0.025	0.850	0.939
Received polio vaccination (3 doses) (Children age 12–23 months)									
Urban	0.937	0.017	311	229	0.013	1.278	0.018	0.902	0.971
Rural	0.889	0.032	252	362	0.020	1.607	0.036	0.826	0.953
Total	0.908	0.021	563	591	0.012	1.738	0.023	0.866	0.949
Received measles vaccination (Children age 12–23 months)									
Urban	0.843	0.028	311	229	0.020	1.391	0.033	0.787	0.898
Rural	0.843	0.033	252	362	0.023	1.423	0.039	0.778	0.909
Total	0.843	0.023	563	591	0.015	1.510	0.027	0.798	0.888
Fully vaccinated (Children age 12–23 months)									
Urban	0.809	0.033	311	229	0.022	1.543	0.041	0.742	0.876
Rural	0.768	0.040	252	362	0.027	1.487	0.052	0.689	0.847
Total	0.784	0.027	563	591	0.017	1.615	0.035	0.729	0.839
Received vitamin A (Children age 12–35 months)									
Urban	0.600	0.036	638	449	0.020	1.812	0.060	0.528	0.672
Rural	0.677	0.037	480	691	0.022	1.716	0.055	0.603	0.751
Total	0.647	0.027	1118	1140	0.014	1.846	0.041	0.593	0.700
Had reproductive health problem (Currently married women age 15–49)									
Urban	0.457	0.025	2951	2044	0.009	2.725	0.055	0.407	0.507
Rural	0.360	0.020	2031	2919	0.011	1.865	0.055	0.320	0.400
Total	0.400	0.016	4982	4963	0.007	2.246	0.039	0.369	0.431
Not involved in any decisionmaking (Ever-married women age 15–49)									
Urban	0.063	0.006	3191	2229	0.004	1.466	0.100	0.050	0.075
Rural	0.078	0.008	2200	3162	0.006	1.417	0.104	0.061	0.094
Total	0.072	0.005	5391	5391	0.004	1.537	0.075	0.061	0.082
Ever beaten or physically mistreated since age 15 (Ever-married women age 15–49)									
Urban	0.169	0.015	3191	2229	0.007	2.289	0.090	0.139	0.200
Rural	0.189	0.011	2200	3162	0.008	1.305	0.058	0.168	0.211
Total	0.181	0.009	5391	5391	0.005	1.687	0.049	0.163	0.199
Not worked in past 12 months (Ever-married women age 15–49)									
Urban	0.690	0.019	3191	2229	0.008	2.262	0.027	0.653	0.727
Rural	0.269	0.021	2200	3162	0.009	2.198	0.077	0.227	0.311
Total	0.443	0.021	5391	5391	0.007	3.155	0.048	0.400	0.486

Table A.2 Sampling errors, Maharashtra, 1999 (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	0.448	0.015	2925	2074	0.009	1.600	0.033	0.418	0.477
Rural	0.512	0.019	2091	2942	0.011	1.745	0.037	0.474	0.550
Total	0.485	0.013	5016	5016	0.007	1.821	0.026	0.460	0.511
Anaemic children (Children age 6–35 months)									
Urban	0.728	0.020	687	489	0.017	1.185	0.028	0.688	0.768
Rural	0.780	0.021	546	769	0.018	1.171	0.027	0.738	0.821
Total	0.760	0.015	1233	1258	0.012	1.242	0.020	0.729	0.790

Table A.2 Sampling errors, Maharashtra, 1999 (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	0.114	0.051	2.015	2.472
Rural	2.740	0.098	0.036	2.545	2.936
Total	2.519	0.073	0.029	2.372	2.665
Age-specific fertility rate (Women age 15–19)					
Urban	0.938	0.009	0.095	0.076	0.112
Rural	0.156	0.010	0.064	0.136	0.176
Total	0.129	0.007	0.053	0.115	0.142
Age-specific fertility rate (Women age 20–24)					
Urban	0.185	0.009	0.050	0.166	0.203
Rural	0.254	0.012	0.048	0.229	0.278
Total	0.223	0.008	0.035	0.207	0.238
Age-specific fertility rate (Women age 25–29)					
Urban	0.111	0.009	0.082	0.092	0.129
Rural	0.101	0.010	0.094	0.082	0.120
Total	0.106	0.007	0.063	0.092	0.119
Age-specific fertility rate (Women age 30–34)					
Urban	0.045	0.007	0.145	0.032	0.058
Rural	0.026	0.006	0.213	0.015	0.038
Total	0.034	0.004	0.124	0.026	0.043
Age-specific fertility rate (Women age 35–39)					
Urban	0.014	0.004	0.261	0.007	0.022
Rural	0.010	0.003	0.320	0.004	0.017
Total	0.012	0.002	0.201	0.007	0.017

Table A.2 Sampling errors, Maharashtra, 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	24.728	5.618	0.227	13.491	35.965
Rural	36.684	6.064	0.165	24.556	48.813
Total	31.998	4.249	0.133	23.499	40.497
Infant mortality $_1q_0$ (5-year period preceding survey)					
Urban	32.974	5.563	0.169	21.847	44.101
Rural	50.632	6.629	0.131	37.374	63.890
Total	43.722	4.566	0.104	34.591	52.853
Child mortality $_4q_1$ (5-year period preceding survey)					
Urban	10.151	3.121	0.307	3.909	16.394
Rural	18.045	3.063	0.170	11.920	24.170
Total	14.985	2.227	0.149	10.531	19.439
Under-five mortality $_5q_0$ (5-year period preceding survey)					
Urban	42.791	6.854	0.160	29.083	56.499
Rural	67.763	6.640	0.098	54.483	81.043
Total	58.052	4.842	0.083	48.369	67.735
Crude death rate (Based on Household Questionnaire)					
Urban	7.891	0.743	0.094	6.404	9.378
Rural	9.950	0.766	0.077	8.418	11.483
Total	9.067	0.559	0.062	7.948	10.185
Crude birth rate (Based on women's birth history)					
Urban	21.623	1.179	0.055	19.264	23.982
Rural	23.819	1.002	0.042	21.815	25.823
Total	22.998	0.772	0.034	21.454	24.541

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. 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 Maharashtra, information on age was missing for only 4 persons out of 29,774 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 Maharashtra.

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 Maharashtra, for age at first marriage, woman's education, and prevalence of diarrhoea in the two weeks preceding the survey, almost complete information is available (Table B.3). Missing information is slightly higher for the age at death of children who died (2 percent), and is still higher for the month of birth of children born in the past 15 years (6 percent). It is important to note, however, that the year of birth is recorded in almost every case in which the month is missing. Data on height and weight of children are available for about 95 percent of living children age 0–35 months, which is acceptable. Some children could not be measured because they were not at

Table B.1 Household age distribution

Single-year age distribution of *de facto* household population by sex (weighted), Maharashtra, 1999

	Male		Female			Male		Female	
Age	Number	Percent	Number	Percent	Age	Number	Percent	Number	Percent
< 1	353	2.3	307	2.1	38	158	1.0	191	1.3
1	278	1.8	313	2.2	39	110	0.7	119	0.8
2	363	2.4	303	2.1	40	409	2.7	265	1.8
3	279	1.8	280	1.9	41	81	0.5	90	0.6
4	358	2.3	309	2.1	42	187	1.2	168	1.2
5	344	2.3	341	2.4	43	74	0.5	88	0.6
6	341	2.2	299	2.1	44	93	0.6	97	0.7
7	297	1.9	338	2.3	45	280	1.8	244	1.7
8	370	2.4	334	2.3	46	97	0.6	107	0.7
9	331	2.2	295	2.0	47	94	0.6	89	0.6
10	460	3.0	372	2.6	48	108	0.7	121	0.8
11	338	2.2	282	1.9	49	96	0.6	21	0.1
12	435	2.8	370	2.6	50	245	1.6	118	0.8
13	324	2.1	329	2.3	51	76	0.5	74	0.5
14	341	2.2	303	2.1	52	98	0.6	118	0.8
15	353	2.3	296	2.0	53	52	0.3	99	0.7
16	338	2.2	337	2.3	54	64	0.4	77	0.5
17	315	2.1	265	1.8	55	155	1.0	197	1.4
18	351	2.3	371	2.6	56	54	0.4	85	0.6
19	241	1.6	247	1.7	57	55	0.4	50	0.3
20	345	2.3	343	2.4	58	86	0.6	73	0.5
21	240	1.6	202	1.4	59	43	0.3	35	0.2
22	353	2.3	340	2.4	60	214	1.4	259	1.8
23	230	1.5	235	1.6	61	29	0.2	32	0.2
24	246	1.6	247	1.7	62	84	0.5	111	0.8
25	377	2.5	382	2.6	63	22	0.1	37	0.3
26	248	1.6	251	1.7	64	42	0.3	47	0.3
27	235	1.5	214	1.5	65	194	1.3	251	1.7
28	255	1.7	259	1.8	66	42	0.3	40	0.3
29	147	1.0	178	1.2	67	33	0.2	27	0.2
30	450	2.9	420	2.9	68	41	0.3	25	0.2
31	124	0.8	152	1.0	69	44	0.3	22	0.2
32	264	1.7	230	1.6	70+	465	3.0	415	2.9
33	110	0.7	116	0.8	Don't know/				
34	186	1.2	167	1.1	missing	3	0.0	1	0.0
35	477	3.1	375	2.6					
36	167	1.1	173	1.2	Total	15,290	100.0	14,484	100.0
37	99	0.6	110	0.8					

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

home or they were ill at the time of the survey. In some cases when the child was at home, either the child refused to be measured or the mother refused to allow the child to be measured. Data on the level of women's haemoglobin are available for 94 percent of respondents and data on the level of children's haemoglobin are available for 87 percent of children. 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. Nevertheless, the response rate for anaemia testing both for women and children is acceptable.

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. Overall, 93 percent of living children listed in the birth history had complete birth dates recorded, as did 74 percent of children who had died. The completeness of data on birth

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), Maharashtra, 1999

Age	All women	Ever-married women	Interviewed women		Percent interviewed
			Number	Percent	
10–14	1,656	21	NA	NA	NA
15–19	1,517	437	398	7.0	91.0
20–24	1,368	1,062	1,008	17.8	94.9
25–29	1,284	1,207	1,152	20.3	95.4
30–34	1,084	1,064	1,012	17.9	95.1
35–39	968	957	914	16.1	95.5
40–44	708	702	660	11.6	94.0
45–49	582	576	520	9.2	90.4
50–54	486	483	NA	NA	NA
15–49	7,512	6,005	5,665	100.0	94.3

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), Maharashtra, 1999

Indicator	Reference group	Percentage missing information	Number of cases
Birth date	Births in past 15 years		
Month only		6.27	9,366
Month and year		0.79	9,366
Age at death	Deaths to births in past 15 years	1.87	676
Age at first marriage	Ever-married women age 15–49	0.05	5,391
Woman's education	Ever-married women age 15–49	0.00	5,391
Anthropometry	Living children age 0–35 months		
Height		5.41	1,736
Weight		5.74	1,736
Height or weight		5.74	1,736
Woman's haemoglobin level	Ever-married women age 15–49	6.07	5,391
Child's haemoglobin level	Living children age 6–35 months	13.17	1,462
Diarrhoea in past 2 weeks	Living children age 1–35 months	0.37	1,704

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), Maharashtra, 1999

	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	201	6	207	100.0	100.0	100.0	983	804	977	NA	NA	NA
1998	575	32	607	99.7	93.4	99.4	920	812	914	NC	NC	NC
1997	590	23	613	99.2	87.8	98.8	1,074	618	1,052	107.0	81.6	105.8
1996	527	25	553	98.6	87.8	98.1	739	1,103	753	90.5	79.5	90.0
1995	575	40	615	96.6	88.5	96.1	915	1,404	940	102.0	108.8	102.4
1994	600	48	648	94.3	75.3	92.9	876	1,069	889	103.1	124.1	104.4
1993	589	38	626	94.3	80.5	93.4	1,036	610	1,005	99.6	76.9	97.9
1992	582	50	631	93.7	88.9	93.3	1,050	684	1,016	103.2	106.8	103.4
1991	539	55	594	93.2	78.7	91.8	935	1,039	944	91.8	108.7	93.2
1990	592	52	644	92.3	72.0	90.6	965	1,574	1,004	102.8	97.6	102.3
1989	613	52	665	94.5	72.7	92.8	903	1,103	917	102.2	108.0	102.7
1988	607	44	651	91.9	75.8	90.8	905	863	902	100.0	78.2	98.2
1993–97	2,881	174	3,055	96.5	82.9	95.8	925	943	926	NA	NA	NA
1988–92	2,931	253	3,185	93.1	77.6	91.9	949	1,022	955	NA	NA	NA
1983–87	2,790	289	3,079	92.0	74.7	90.4	889	809	881	NA	NA	NA
1978–82	2,164	318	2,483	91.1	74.4	89.0	907	843	899	NA	NA	NA
1977 or earlier	1,800	460	2,260	88.9	66.6	84.4	945	1,133	981	NA	NA	NA
All	13,342	1,533	14,876	93.1	74.1	91.2	923	955	926	NA	NA	NA

NA: Not applicable
NC: Not calculated because full-year data were not collected for 1999
[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

dates for surviving children is good overall and is excellent in recent years. The completeness of data on birth dates for nonsurviving children is less satisfactory overall, but is again better in recent years. 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 January 1996, 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, the sharp drop in the annual number of births in the year 1996 (for both surviving and 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,

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), Maharashtra, 1999

	Years preceding survey			
Age at death (days)	0–4	5–9	10–14	0–14
< 1	14	38	28	81
1	25	30	19	75
2	8	16	10	34
3	8	10	9	27
4	5	6	4	15
5	10	7	4	21
6	4	2	3	9
7	3	3	5	10
8	1	6	3	10
9	1	0	0	1
10	2	3	3	7
11	0	1	1	2
12	0	1	6	8
13	1	1	0	3
14	2	0	0	2
15	6	9	16	30
16	0	0	3	3
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	3	3	0	5
21	0	4	0	5
22	0	0	0	0
23	0	0	0	0
24	0	2	0	2
25	0	0	0	1
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	94	141	116	351
Percent early neonatal[1]	79.2	76.4	68.2	74.4

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

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 Maharashtra, because the ratios of deaths under seven days to all neonatal deaths are consistently high (a value of less than 25 percent is often used as a guideline to indicate underreporting of early neonatal deaths). The estimates are 79 percent for 0–4 years, 76 for percent for 5–9 years, and 68 percent 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 5, 10, 12, 15, 20, and 21 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

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), Maharashtra, 1999

	Years preceding survey			
Age at death (months)	0–4	5–9	10–14	0–14
< 1	94	141	116	351
1	4	14	14	32
2	3	8	13	24
3	6	4	5	14
4	6	3	1	10
5	2	6	6	14
6	1	7	2	11
7	5	4	3	12
8	1	3	0	4
9	0	5	7	12
10	0	1	3	5
11	1	4	2	7
12	3	3	8	15
13	0	0	2	2
14	0	1	0	1
15	0	1	0	2
16	0	0	0	0
17	1	0	0	1
18	1	6	4	12
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	0	0	0	0
1 year	3	8	13	24
0–11 months	125	200	172	497
Percent neonatal[1]	75.3	70.6	67.6	70.7

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

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 year or older. Thus, heaping at 12 months can bias the mortality estimates because a certain fraction of these deaths, which are reported to have occurred after infancy 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 the child mortality rate 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 that there is substantial heaping of deaths at 5, 9, 12, and 18 months of age. The heaping in the past five years is less evident. The degree of digit preference in Maharashtra is not serious enough to alter substantially the estimated mortality rates, probably due to the strong emphasis on this potential problem during training of the interviewers.