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, Tamil Nadu, 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, Tamil Nadu, 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	1005	16.773	5150	3852	15.236	1.101	0.017	971	1039
Rural	1049	16.313	5820	7009	14.508	1.124	0.016	1017	1082
Total	1034	12.106	10970	10860	10.565	1.146	0.012	1009	1058
Illiterate (*De facto* household population age 6 and above)									
Urban	0.187	0.017	9024	6806	0.005	3.299	0.090	0.153	0.221
Rural	0.380	0.013	10489	12630	0.006	2.314	0.034	0.354	0.406
Total	0.312	0.013	19513	19436	0.004	3.202	0.042	0.286	0.339
Have tuberculosis (1,000 *de jure* household population)									
Urban	4.307	0.974	10215	7705	0.704	1.383	0.226	2.359	6.254
Rural	5.047	0.798	12064	14527	0.670	1.190	0.158	3.452	6.643
Total	4.791	0.621	22279	22232	0.488	1.271	0.130	3.550	6.032
Salt iodized at 15 ppm or more (Households)									
Large city	0.355	0.039	1110	347	0.014	2.719	0.110	0.277	0.433
Small city	0.384	0.050	524	594	0.021	2.372	0.131	0.283	0.484
Town	0.395	0.060	754	856	0.018	3.380	0.152	0.275	0.516
Rural	0.124	0.014	2893	3484	0.006	2.242	0.111	0.096	0.151
Total	0.212	0.018	5281	5281	0.006	3.224	0.085	0.176	0.248
Illiterate (Ever-married women age 15–49)									
Urban	0.278	0.028	2113	1620	0.010	2.905	0.102	0.222	0.335
Rural	0.579	0.019	2563	3056	0.010	1.931	0.033	0.541	0.617
Total	0.475	0.020	4676	4676	0.007	2.795	0.043	0.434	0.516
High school complete and above (Ever-married women age 15–49)									
Urban	0.284	0.027	2113	1620	0.010	2.721	0.094	0.231	0.337
Rural	0.092	0.008	2563	3056	0.006	1.354	0.084	0.077	0.108
Total	0.159	0.013	4676	4676	0.005	2.473	0.083	0.132	0.185
Currently married women (Ever-married women age 15–49)									
Urban	0.924	0.006	2113	1620	0.006	1.113	0.007	0.911	0.937
Rural	0.899	0.006	2563	3056	0.006	1.093	0.007	0.886	0.912
Total	0.908	0.005	4676	4676	0.004	1.160	0.005	0.898	0.918
Number of children ever born (Currently married women age 15–49)									
Urban	2.200	0.043	1961	1497	0.032	1.320	0.019	2.114	2.285
Rural	2.526	0.047	2305	2748	0.035	1.346	0.019	2.432	2.621
Total	2.411	0.037	4266	4245	0.025	1.500	0.015	2.337	2.485
Number of living children (Currently married women age 15–49)									
Urban	2.037	0.039	1961	1497	0.029	1.357	0.019	1.958	2.116
Rural	2.201	0.039	2305	2748	0.029	1.361	0.018	2.123	2.279
Total	2.143	0.030	4266	4245	0.021	1.443	0.014	2.084	2.203
Have ever used any method (Currently married women age 15–49)									
Urban	0.647	0.018	1961	1497	0.011	1.639	0.027	0.612	0.683
Rural	0.525	0.013	2305	2748	0.010	1.282	0.025	0.498	0.552
Total	0.568	0.012	4266	4245	0.008	1.543	0.021	0.545	0.592

Table A.2 Sampling errors, Tamil Nadu, 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	0.582	0.017	1961	1497	0.011	1.490	0.029	0.549	0.615
Rural	0.488	0.012	2305	2748	0.010	1.199	0.026	0.463	0.513
Total	0.521	0.010	4266	4245	0.008	1.372	0.020	0.500	0.542
Currently using any modern method (Currently married women age 15–49)									
Urban	0.551	0.016	1961	1497	0.011	1.443	0.029	0.519	0.584
Rural	0.476	0.013	2305	2748	0.010	1.244	0.027	0.450	0.502
Total	0.503	0.010	4266	4245	0.008	1.367	0.021	0.482	0.524
Currently using pills (Currently married women age 15–49)									
Urban	0.004	0.002	1961	1497	0.001	1.281	0.437	0.001	0.008
Rural	0.003	0.001	2305	2748	0.001	1.120	0.457	0.000	0.005
Total	0.003	0.001	4266	4245	0.001	1.174	0.316	0.001	0.005
Currently using IUD (Currently married women age 15–49)									
Urban	0.050	0.007	1961	1497	0.005	1.446	0.142	0.036	0.064
Rural	0.011	0.003	2305	2748	0.002	1.230	0.245	0.006	0.016
Total	0.025	0.004	4266	4245	0.002	1.498	0.144	0.018	0.032
Currently using condoms (Currently married women age 15–49)									
Urban	0.031	0.004	1961	1497	0.004	1.150	0.146	0.022	0.040
Rural	0.007	0.002	2305	2748	0.002	1.110	0.277	0.003	0.011
Total	0.015	0.002	4266	4245	0.002	1.198	0.147	0.011	0.020
Currently using female sterilization (Currently married women age 15–49)									
Urban	0.460	0.017	1961	1497	0.011	1.506	0.037	0.426	0.494
Rural	0.447	0.013	2305	2748	0.010	1.292	0.030	0.420	0.474
Total	0.452	0.010	4266	4245	0.008	1.373	0.023	0.431	0.473
Currently using male sterilization (Currently married women age 15–49)									
Urban	0.006	0.002	1961	1497	0.002	1.038	0.308	0.002	0.009
Rural	0.009	0.003	2305	2748	0.002	1.650	0.367	0.002	0.015
Total	0.008	0.002	4266	4245	0.001	1.615	0.282	0.003	0.012
Currently using rhythm/safe period (Currently married women age 15–49)									
Urban	0.022	0.004	1961	1497	0.003	1.174	0.179	0.014	0.029
Rural	0.008	0.002	2305	2748	0.002	1.233	0.290	0.003	0.012
Total	0.013	0.002	4266	4245	0.002	1.183	0.160	0.009	0.017
Using public source for modern method (Current users of modern methods)									
Urban	0.650	0.030	1110	825	0.014	2.061	0.045	0.591	0.709
Rural	0.789	0.024	1098	1309	0.012	1.933	0.030	0.742	0.837
Total	0.735	0.020	2208	2134	0.009	2.079	0.027	0.696	0.775
Do not want any more children (Currently married women age 15–49)									
Urban	0.252	0.014	1961	1497	0.010	1.394	0.054	0.225	0.280
Rural	0.217	0.012	2305	2748	0.009	1.431	0.057	0.193	0.242
Total	0.230	0.009	4266	4245	0.006	1.463	0.041	0.211	0.249
Want to delay birth at least two years (Currently married women age 15–49)									
Urban	0.107	0.007	1961	1497	0.007	1.044	0.068	0.092	0.122
Rural	0.124	0.007	2305	2748	0.007	0.988	0.055	0.111	0.138
Total	0.118	0.005	4266	4245	0.005	1.039	0.043	0.108	0.128

Table A.2 Sampling errors, Tamil Nadu, 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	1.997	0.027	2052	1567	0.012	2.176	0.014	1.943	2.052
Rural	2.063	0.021	2482	2959	0.011	1.952	0.010	2.022	2.105
Total	2.040	0.017	4534	4526	0.008	2.055	0.008	2.007	2.074
Ideal number of sons (Ever-married women age 15–49)									
Urban	0.682	0.028	2052	1567	0.014	2.064	0.041	0.626	0.738
Rural	0.799	0.019	2482	2959	0.013	1.483	0.023	0.762	0.836
Total	0.758	0.016	4534	4526	0.009	1.750	0.021	0.726	0.791
Ideal number of daughters (Ever-married women age 15–49)									
Urban	0.622	0.025	2052	1567	0.012	2.041	0.040	0.573	0.671
Rural	0.686	0.013	2482	2959	0.010	1.284	0.019	0.660	0.713
Total	0.664	0.012	4534	4526	0.008	1.583	0.019	0.639	0.689
Visited by health/family planning worker (Ever-married women age 15–49)									
Urban	0.211	0.020	2113	1620	0.009	2.294	0.097	0.170	0.251
Rural	0.287	0.015	2563	3056	0.009	1.669	0.052	0.257	0.316
Total	0.260	0.012	4676	4676	0.006	1.927	0.048	0.236	0.285
Received no antenatal check-up (Births in past 3 years)									
Urban	0.008	0.005	597	467	0.004	1.483	0.677	0.000	0.018
Rural	0.013	0.004	748	892	0.004	1.014	0.318	0.005	0.022
Total	0.012	0.003	1345	1359	0.003	1.153	0.290	0.005	0.018
Received iron and folic acid tablets or syrup (Births in past 3 years)									
Urban	0.929	0.018	597	467	0.010	1.722	0.019	0.893	0.966
Rural	0.933	0.013	748	892	0.009	1.377	0.014	0.908	0.958
Total	0.932	0.010	1345	1359	0.007	1.500	0.011	0.911	0.952
Received medical assistance during delivery (Births in past 3 years)									
Urban	0.949	0.026	597	467	0.009	2.769	0.027	0.898	1.000
Rural	0.779	0.025	748	892	0.016	1.537	0.032	0.730	0.828
Total	0.838	0.020	1345	1359	0.011	1.863	0.023	0.798	0.877
Received postpartum check-up (Noninstitutional births in past 3 years)									
Urban	0.638	0.139	34	35	0.084	1.656	0.217	0.361	0.915
Rural	0.514	0.041	206	246	0.035	1.163	0.079	0.433	0.596
Total	0.530	0.037	240	281	0.032	1.151	0.070	0.456	0.604
Had diarrhoea in the past 2 weeks (Children under 3 years)									
Urban	0.150	0.020	568	445	0.015	1.328	0.133	0.110	0.190
Rural	0.140	0.015	707	843	0.013	1.115	0.104	0.111	0.169
Total	0.144	0.012	1275	1288	0.010	1.191	0.081	0.120	0.167
Treated with ORS packets (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.290	0.053	87	67	0.049	1.096	0.184	0.183	0.397
Rural	0.273	0.046	99	118	0.046	1.012	0.170	0.180	0.366
Total	0.279	0.035	186	185	0.034	1.041	0.125	0.210	0.349
Taken to a health facility/provider for diarrhoea (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.723	0.044	87	67	0.049	0.898	0.060	0.636	0.810
Rural	0.645	0.050	99	118	0.050	1.015	0.078	0.545	0.746
Total	0.673	0.036	186	185	0.035	1.006	0.053	0.602	0.745

Table A.2 Sampling errors, Tamil Nadu, 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	0.554	0.040	197	151	0.035	1.127	0.072	0.474	0.633
Rural	0.408	0.035	240	286	0.032	1.098	0.086	0.338	0.478
Total	0.458	0.028	437	438	0.024	1.154	0.060	0.403	0.513
Received BCG vaccination (Children age 12–23 months)									
Urban	1.000	0.000	197	151	0.000	NC	NC	1.000	1.000
Rural	0.979	0.009	240	286	0.009	0.975	0.009	0.961	0.997
Total	0.986	0.006	437	438	0.006	1.075	0.006	0.974	0.998
Received DPT vaccination (3 doses) (Children age 12–23 months)									
Urban	0.998	0.002	197	151	0.003	0.707	0.002	0.993	1.000
Rural	0.950	0.014	240	286	0.014	1.001	0.015	0.922	0.978
Total	0.967	0.010	437	438	0.009	1.114	0.010	0.947	0.986
Received polio vaccination (3 doses) (Children age 12–23 months)									
Urban	0.998	0.002	197	151	0.003	0.707	0.002	0.993	1.000
Rural	0.971	0.013	240	286	0.011	1.179	0.013	0.945	0.997
Total	0.980	0.009	437	438	0.007	1.297	0.009	0.963	0.998
Received measles vaccination (Children age 12–23 months)									
Urban	0.968	0.015	197	151	0.012	1.167	0.015	0.939	0.997
Rural	0.867	0.021	240	286	0.022	0.970	0.025	0.824	0.909
Total	0.902	0.016	437	438	0.014	1.112	0.018	0.870	0.933
Fully vaccinated (Children age 12–23 months)									
Urban	0.968	0.015	197	151	0.012	1.167	0.015	0.939	0.997
Rural	0.846	0.024	240	286	0.023	1.020	0.028	0.798	0.893
Total	0.888	0.018	437	438	0.015	1.172	0.020	0.853	0.923
Received vitamin A (Children age 12–35 months)									
Urban	0.193	0.025	399	316	0.020	1.248	0.129	0.143	0.243
Rural	0.145	0.020	470	560	0.017	1.218	0.140	0.104	0.185
Total	0.162	0.016	869	876	0.013	1.243	0.098	0.130	0.194
Had reproductive health problem (Currently married women age 15–49)									
Urban	0.279	0.015	1961	1497	0.010	1.521	0.055	0.248	0.310
Rural	0.278	0.012	2305	2748	0.009	1.308	0.044	0.254	0.303
Total	0.278	0.010	4266	4245	0.007	1.389	0.034	0.259	0.298
Not involved in any decisionmaking (Ever-married women age 15–49)									
Urban	0.029	0.005	2113	1620	0.004	1.265	0.158	0.020	0.039
Rural	0.021	0.003	2563	3056	0.003	1.042	0.140	0.015	0.027
Total	0.024	0.003	4676	4676	0.002	1.121	0.105	0.019	0.029
Ever beaten or physically mistreated since age 15 (Ever-married women age 15–49)									
Urban	0.356	0.014	2113	1620	0.010	1.342	0.039	0.328	0.384
Rural	0.429	0.015	2563	3056	0.010	1.494	0.034	0.400	0.458
Total	0.404	0.011	4676	4676	0.007	1.546	0.027	0.381	0.426
Not worked in past 12 months (Ever-married women age 15–49)									
Urban	0.627	0.031	2113	1620	0.011	2.993	0.050	0.564	0.690
Rural	0.374	0.019	2563	3056	0.010	2.018	0.052	0.336	0.413
Total	0.462	0.019	4676	4676	0.007	2.640	0.042	0.423	0.500

Table A.2 Sampling errors, Tamil Nadu, 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	0.516	0.023	2086	1591	0.011	2.069	0.044	0.471	0.561
Rural	0.591	0.019	2505	3000	0.010	1.962	0.033	0.552	0.629
Total	0.565	0.015	4591	4591	0.007	2.095	0.027	0.534	0.595
Anaemic children (Children age 6–35 months)									
Urban	0.662	0.030	475	366	0.022	1.358	0.045	0.603	0.721
Rural	0.705	0.020	574	687	0.019	1.075	0.029	0.664	0.746
Total	0.690	0.017	1049	1053	0.014	1.177	0.024	0.657	0.724

Table A.2 Sampling errors, Tamil Nadu, 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	2.107	0.102	0.049	1.903	2.312
Rural	2.230	0.075	0.034	2.081	2.380
Total	2.188	0.060	0.027	2.068	2.308
Age-specific fertility rate (Women age 15–19)					
Urban	0.071	0.011	0.160	0.048	0.093
Rural	0.090	0.007	0.080	0.076	0.104
Total	0.083	0.006	0.071	0.071	0.095
Age-specific fertility rate (Women age 20–24)					
Urban	0.172	0.008	0.048	0.156	0.189
Rural	0.199	0.008	0.042	0.182	0.215
Total	0.189	0.006	0.032	0.177	0.201
Age-specific fertility rate (Women age 25–29)					
Urban	0.122	0.010	0.085	0.102	0.143
Rural	0.120	0.009	0.072	0.103	0.138
Total	0.121	0.007	0.055	0.108	0.134
Age-specific fertility rate (Women age 30–34)					
Urban	0.042	0.008	0.198	0.025	0.058
Rural	0.026	0.005	0.193	0.016	0.036
Total	0.032	0.004	0.137	0.023	0.041
Age-specific fertility rate (Women age 35–39)					
Urban	0.011	0.004	0.323	0.004	0.018
Rural	0.009	0.003	0.359	0.003	0.016
Total	0.010	0.002	0.251	0.005	0.015
Age-specific fertility rate (Women age 40–44)					
Urban	0.004	0.003	0.710	0.000	0.009
Rural	0.002	0.001	0.704	0.000	0.005
Total	0.003	0.001	0.495	0.026	0.005

Table A.2 Sampling errors, Tamil Nadu, 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	28.451	6.341	0.223	15.769	41.134
Rural	38.144	6.157	0.161	25.831	50.458
Total	34.824	4.577	0.131	25.670	43.977
Infant mortality $_1q_0$ (5-year period preceding survey)					
Urban	40.640	6.833	0.168	26.974	54.307
Rural	52.132	6.688	0.128	38.757	65.507
Total	48.168	4.973	0.103	38.221	58.114
Child mortality $_4q_1$ (5-year period preceding survey)					
Urban	9.394	5.072	0.540	0.000	19.537
Rural	19.298	3.739	0.194	11.820	26.776
Total	15.921	3.027	0.190	9.868	21.974
Under-five mortality $_5q_0$ (5-year period preceding survey)					
Urban	49.652	8.609	0.173	32.433	66.871
Rural	70.424	7.734	0.110	54.956	85.893
Total	63.322	5.911	0.093	51.499	75.144
Crude death rate (Based on Household Questionnaire)					
Urban	8.066	0.862	0.107	6.342	9.789
Rural	12.157	0.800	0.066	10.556	13.758
Total	10.739	0.636	0.059	9.467	12.011
Crude birth rate (Based on women's birth history)					
Urban	21.257	1.095	0.052	19.067	23.446
Rural	21.506	0.858	0.040	19.791	23.222
Total	21.410	0.677	0.032	20.055	22.764

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. 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 Tamil Nadu, information on age was missing for only 2 persons out of 22,085 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 Tamil Nadu.

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 Tamil Nadu, the extent of missing information is very low for age at death, age at first marriage, woman's education, and prevalence of diarrhoea in the two weeks preceding the survey (Table B.3). Month of birth was missing for 2 percent of children; however, the year is reported in almost every case in which the month is missing. Data on height and weight of children are available for 96 percent of children under three years of age. Many children could not be measured because they were not at 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

Table B.1 Household age distribution

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

	Male		Female			Male		Female	
Age	Number	Percent	Number	Percent	Age	Number	Percent	Number	Percent
< 1	239	2.2	220	2.0	38	135	1.2	159	1.4
1	234	2.2	207	1.8	39	128	1.2	137	1.2
2	230	2.1	229	2.0	40	228	2.1	168	1.5
3	219	2.0	228	2.0	41	75	0.7	128	1.1
4	221	2.0	191	1.7	42	107	1.0	129	1.1
5	209	1.9	223	2.0	43	78	0.7	108	1.0
6	181	1.7	205	1.8	44	104	1.0	99	0.9
7	247	2.3	208	1.9	45	231	2.1	130	1.2
8	227	2.1	213	1.9	46	91	0.8	99	0.9
9	204	1.9	222	2.0	47	85	0.8	102	0.9
10	229	2.1	194	1.7	48	95	0.9	98	0.9
11	215	2.0	204	1.8	49	105	1.0	110	1.0
12	237	2.2	203	1.8	50	138	1.3	53	0.5
13	189	1.7	225	2.0	51	60	0.6	62	0.6
14	230	2.1	197	1.8	52	83	0.8	98	0.9
15	206	1.9	220	2.0	53	43	0.4	70	0.6
16	184	1.7	211	1.9	54	57	0.5	97	0.9
17	177	1.6	201	1.8	55	133	1.2	203	1.8
18	247	2.3	218	1.9	56	53	0.5	121	1.1
19	197	1.8	210	1.9	57	36	0.3	79	0.7
20	193	1.8	246	2.2	58	83	0.8	104	0.9
21	165	1.5	218	1.9	59	65	0.6	61	0.5
22	196	1.8	254	2.3	60	163	1.5	177	1.6
23	162	1.5	243	2.2	61	61	0.6	49	0.4
24	172	1.6	202	1.8	62	76	0.7	66	0.6
25	236	2.2	234	2.1	63	41	0.4	47	0.4
26	211	1.9	203	1.8	64	31	0.3	48	0.4
27	188	1.7	217	1.9	65	147	1.4	166	1.5
28	197	1.8	216	1.9	66	19	0.2	25	0.2
29	182	1.7	192	1.7	67	31	0.3	33	0.3
30	252	2.3	232	2.1	68	41	0.4	41	0.4
31	114	1.1	149	1.3	69	29	0.3	37	0.3
32	198	1.8	167	1.5	70+	404	3.7	353	3.1
33	126	1.2	160	1.4	Don't				
34	132	1.2	157	1.4	know/				
35	299	2.8	182	1.6	missing	0	0.0	2	0.0
36	129	1.2	142	1.3					
37	129	1.2	126	1.1	Total	10,860	100.0	11,225	100.0

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

be measured. Data on the haemoglobin level of women are available for 98 percent of respondents and data on children's haemoglobin level are available for 95 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.

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, 97 percent of living children listed in the birth history had complete birth dates recorded, as did 92 percent of children who had died. The completeness of data on birth dates for surviving children is very good overall and excellent in recent years. The completeness

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), Tamil Nadu, 1999

Age	All women	Ever-married women	Interviewed women		Percent interviewed
			Number	Percent	
10–14	1,023	2	NA	NA	NA
15–19	1,061	251	250	5.3	99.5
20–24	1,162	786	784	16.6	99.7
25–29	1,060	972	970	20.6	99.8
30–34	865	845	844	17.9	99.8
35–39	746	728	727	15.4	99.8
40–44	632	620	617	13.1	99.6
45–49	539	529	528	11.2	99.8
50–54	380	378	NA	NA	NA
15–49	6,065	4,732	4,720	100.0	99.7

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), Tamil Nadu, 1999

Indicator	Reference group	Percentage missing information	Number of cases
Birth date	Births in past 15 years		
Month only		1.72	6,641
Month and year		0.29	6,641
Age at death	Deaths to births in past 15 years	0.48	492
Age at first marriage	Ever-married women age 15–49	0.03	4,676
Woman's education	Ever-married women age 15–49	0.03	4,676
Anthropometry	Living children age 0–35 months		
Height		3.67	1,308
Weight		3.75	1,308
Height or weight		3.75	1,308
Woman's haemoglobin level	Ever-married women age 15–49	1.92	4,676
Child's haemoglobin level	Living children age 6–35 months	5.33	1,113
Diarrhoea in past 2 weeks	Living children age 1–35 months	0.36	1,288

for nonsurviving children is less satisfactory overall, but is also excellent in recent years. 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 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. The number of births is

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), Tamil Nadu, 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	132	2	134	100.0	100.0	100.0	1,049	0	1,014	NA	NA	NA
1998	449	22	471	100.0	98.5	99.9	865	1,470	886	NC	NC	NC
1997	437	22	460	100.0	98.6	99.9	911	1,375	929	98.3	126.5	99.3
1996	441	13	454	99.7	100.0	99.7	1,003	595	988	101.7	52.2	98.9
1995	430	29	459	99.7	100.0	99.7	921	1,763	958	102.5	104.0	102.6
1994	398	42	440	98.7	96.4	98.4	1,049	792	1,022	94.9	125.6	97.2
1993	409	38	447	99.2	96.1	98.9	1,085	1,473	1,114	101.9	99.3	101.7
1992	405	35	440	99.2	95.7	98.9	862	799	857	97.2	109.5	98.1
1991	423	26	449	98.6	98.8	98.6	989	1,103	995	103.1	90.2	102.2
1990	417	22	438	98.3	94.7	98.1	1,043	396	997	101.9	67.2	99.4
1989	394	39	434	98.0	100.0	98.1	875	751	863	98.3	151.0	101.5
1988	386	30	416	97.8	100.0	98.0	994	882	986	98.8	65.2	95.3
1993–97	2,116	144	2,261	99.5	97.7	99.4	990	1,156	1,000	NA	NA	NA
1988–92	2,025	151	2,176	98.4	98.1	98.4	951	772	937	NA	NA	NA
1983–87	1,928	244	2,172	96.9	90.3	96.1	958	1,067	970	NA	NA	NA
1978–82	1,679	280	1,959	95.9	90.2	95.1	904	907	905	NA	NA	NA
1977 or earlier	1,687	468	2,155	93.6	88.8	92.6	984	752	929	NA	NA	NA
All	10,016	1,312	11,328	97.2	91.6	96.6	955	886	947	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

fairly constant from 1991 to 1999, so there is no indication of omission or displacement of births in 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 successful in this regard in Tamil Nadu.

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 Tamil Nadu, 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

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), Tamil Nadu, 1999

	Years preceding survey			
Age at death (days)	0–4	5–9	10–14	0–14
< 1	22	21	20	63
1	10	11	14	35
2	11	8	10	28
3	9	12	13	34
4	3	2	4	9
5	3	3	2	8
6	5	0	0	5
7	5	0	1	6
8	1	5	2	9
9	1	1	1	4
10	0	6	2	8
11	1	0	1	2
12	0	2	0	2
13	0	0	0	0
14	0	1	1	2
15	0	5	3	7
16	0	0	0	0
17	0	0	1	1
18	2	0	0	2
19	0	0	0	0
20	0	1	4	5
21	0	0	0	0
22	1	1	1	4
23	0	0	0	0
24	0	0	0	0
25	0	2	0	2
26	0	0	0	0
27	0	0	0	0
28	1	1	0	2
29	0	0	0	0
30	1	1	0	2
0–30	77	83	80	240
Percent early neonatal[1]	81.8	67.9	77.4	75.5

[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 the percentage of infant deaths reported to occur at age under one month, for births occurring during five-year periods preceding the survey (weighted), Tamil Nadu, 1999

	Years preceding survey			
Age at death (months)	0–4	5–9	10–14	0–14
< 1	77	83	80	240
1	1	10	13	24
2	2	2	0	4
3	3	1	13	17
4	1	2	5	9
5	1	1	2	5
6	4	4	8	16
7	4	4	3	10
8	1	5	6	12
9	4	1	6	11
10	4	4	6	13
11	0	1	1	2
12	0	1	5	7
13	0	0	4	4
14	0	0	1	1
15	1	1	0	2
16	1	0	0	2
17	0	0	0	0
18	1	6	4	11
19	1	0	0	1
20	1	0	1	2
21	0	0	1	1
22	0	0	0	0
23	0	0	0	0
1 year	0	3	10	13
0–11 months	102	118	142	362
Percent neonatal[1]	75.5	70.6	56.2	66.3

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

deaths). The ratios are 82 for 0–4 years, 68 for 5–9 years, and 77 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, 8, 10, 15, and 20 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 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 there is heaping of deaths at 3, 6, 10,12, and 18 months of age. Digit preference appears not to be serious enough to alter substantially the mortality rates calculated here. Because the extent of heaping on 12 months is minor, probably due to strong emphasis on this potential problem during training of interviewers, adjustment of the infant and child mortality rates is unnecessary.

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