

CHAPTER 3

Methodology

Introduction

This chapter describes the methods and materials used for the study like objectives, theoretical framework, hypotheses, area covered, source of data, sample design, population, unit of analysis, different approaches to analyze the data, brief description of the multivariate techniques used, quality of data and limitations of the study.

Objectives of the study

The principle objective of the study is to understand the role of social determinants, particularly social group affiliation on immunization coverage among children aged 12-59 months and laying emphasis on SCs and STs in Odisha, Rajasthan and Madhya Pradesh by using the data from the NFHS-3. It is confined to these three major EAG states, because of their diversity in terms of social group composition, different ecological backdrop and poor health status of children (infant and child mortality rates are higher than the national average) and coverage of EPI vaccines lowest to medium. The intermediate objectives of the study, leading to the principal objectives, are as follows:

- To examine the differentials in coverage of different vaccines and patterns of partial vaccination (drop-out, left-out and missed-out) by source of vaccination data and across social groups;
- To understand the effect of social determinants on child immunization, particularly the role of caste/tribe sequentially controlling for other risk factors;
- To understand whether social group affiliation and programme factors interact with each other in predicting child immunization; and,
- To understand the effects of coverage of full immunization and social determinants, particularly the role of social group affiliation (caste/tribe) on Vitamin A supplementation coverage, sequentially controlling for other risk factors.

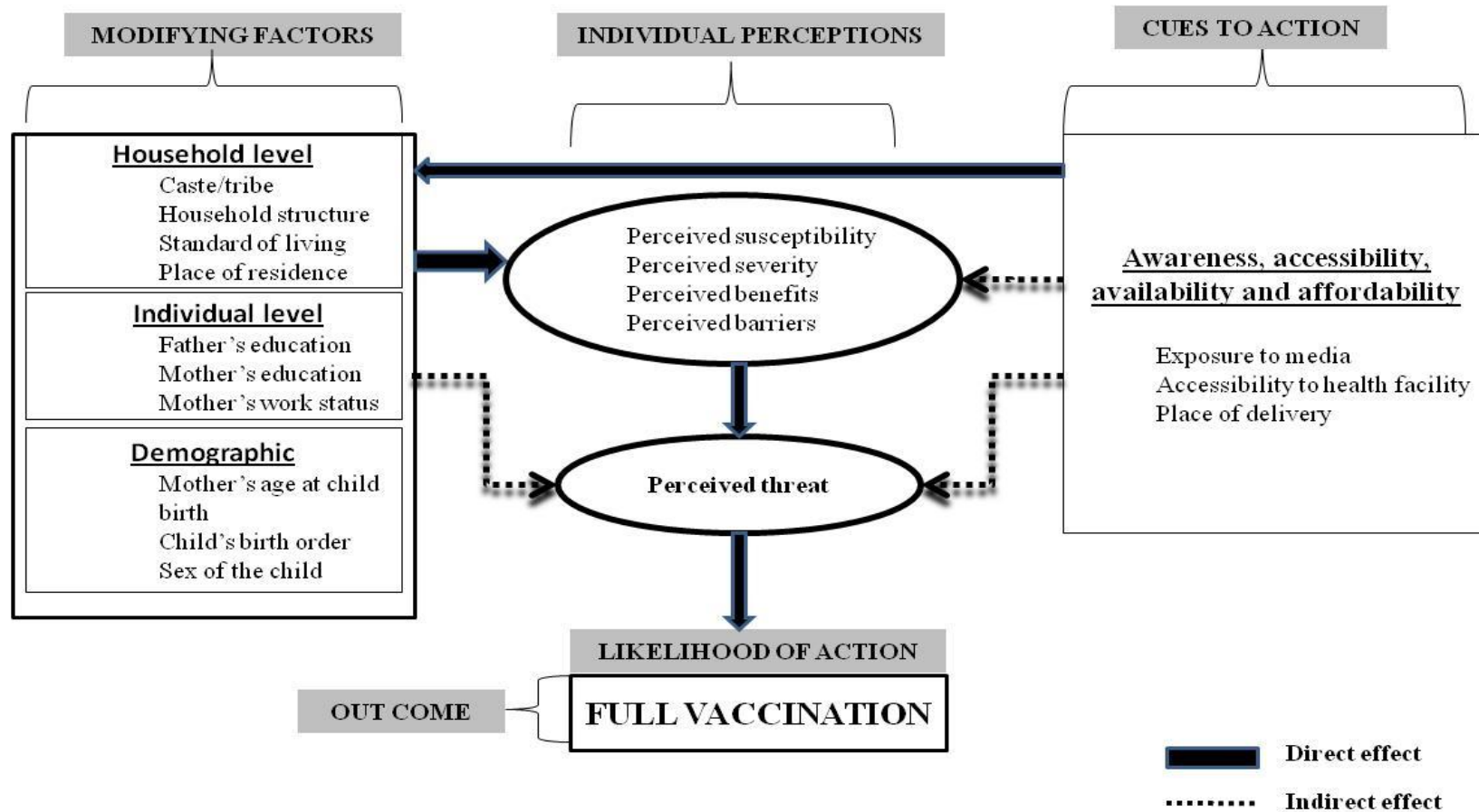
Conceptual framework of the study

This study has been conceptualized based on two major conceptual models - Health Belief Model (Rosenstock, 1990) and Tanahashi model for health services coverage (Tanahashi, 1978) to address the social group inequities in childhood immunization coverage using the data from the Third National Family Health Survey (NFHS-3) in Rajasthan, Odisha and Madhya Pradesh. Both of the models primarily contain three aspects of health services coverage, modifying factors – individual perception and cues to action factors (related to coverage, availability, accessibility, acceptability and contact). Schematically the conceptual model of this study has been presented in Figure 1, and detailed description is given in the following paragraphs.

The Health Belief Model provides a set of constructs for identifying and understanding the multiple factors that influence the demand and delivery of a vaccine. It is an influential and widely used theoretical model initially developed by the U.S. Public Health Service (Rosenstock, 1974; Janz Champion and Strecher, 2002). In the Health Belief Model, there are five key components which includes: First, the perception of threat, which is conceived further as two components: perceived severity and susceptibility to an adverse outcome; Second, perceived outcome expectations, which are examined as perceived benefits and perceived barriers to performing a protective behaviour; Third, modifiers, which influence the individual's response to the model, such as socio-demographic factors and cultural beliefs. Modifiers are demographic, social or psychological in nature, which are likely to influence the health outcome resulting from some action.

The Fourth component of this model is likelihood of the behavioural outcome. An important tenet of the Health Belief Model is the idea of perceived threat which is the combination of an individual's perception of severity of a health problem and that individual's perceived susceptibility of being affected by a potential health risk. Health beliefs and behaviour are the perception of outcome expectations — what a person feels will be the result of some action — considered as either perceived benefits or perceived barriers to achieving a desired outcome. Perceived outcomes of expectations are what persons feel to gain benefits from behavioural change or feel a negative impact of barriers from such behavioural change.

Figure 1: Conceptual framework



Source: Rosenstock, 1990 and Tanahashi, 1978

The final ingredient (*fifth construct*) in the Health Belief Model is a stimulus or cue to action factors. These factors promptly influence a person to initiate the completion of a recommended behaviour change or action. It is a facilitative mechanism and context for introducing a protective behaviour (Figure 1).

In Health Belief Model, however, it was not determined what the cues to actions and variables are. Therefore, at this point Tanahashi proposed concept of health service coverage model (Tanahashi, 1978) is introduced in this study, which expresses the extent of interaction between the health services and the people for whom it is intended. This model distinguishes four different levels of healthcare coverage viz., (i) Availability; (ii) Accessibility; (iii) Acceptability; and (iv) Contact.

In the context of the child healthcare seeking, a number of factors may affect the availability, accessibility, acceptability and contact coverage of health services. They are summarized as follows. (a) *Availability coverage* tackles the issue of the health care resources being available and for whom such resources are available. (b) *Accessibility coverage* assesses how readily accessible the health resources are and for whom they are accessible to. This term is sometimes used to represent physical or financial barriers to access. *Acceptability coverage* simply asks if the resources or the service is acceptable to the intended population and for which population is the intervention for. This includes social, cultural and perception and financial barriers to using services. (d) *Contact coverage* is the intended population making contact with the intervention or utilization.

Major limitations of the Health Belief Model includes (a) most HBM-based research has incorporated only selected components of the HBM, thereby not testing the model as a whole; (b) as a psychological model it does not take into consideration other factors, such as environmental or economic factors, that may influence health behaviours. Therefore, based on the availability of data and considering the limitations of the HBM, analytical framework formulated for this study in the following lines.

In this study, social group affiliation is the primary variable of interest in the analysis. Based on the availability of data in the NFHS-3, following variables were considered as modifiers, viz., place of residence, household structure, household's standard of living (index), father's education, mother's education, mother's work status, mother's age at the time of child birth, birth order and sex of the child (socio-economic

and demographic characteristics at household, parental and child level). The literature indicates that these factors are significantly associated with seeking and using skilled care (uptake of childhood immunization).

Due to lack of more information on availability and quality of services in the NFHS-3 data, this study has focused solely on three ‘demand side’ cues to action variables (programme exposure variables such as exposure to electronic media, place of delivery and mother’s access to health facility). The justification for inclusion of the above three cues to action variables is obvious and needs no explanation. However, the rationale may be provided in brief.

The attitude within a household could be changed through interventions by the mode of communication, if it is interpersonal or exposure to mass media. If the parent or caregiver of the child is exposed to electronic or print media then the awareness and attitude must be changed for child immunization, because the Government of India has been broadcasting the necessary information about the issue since its inception through electronic and print media. Place of delivery is another important factor that influences the coverage of childhood immunization. It not only indicates the availability of the facility but also its accessibility and affordability to receive health care. Apart from this, if a woman delivers in an institutional facility, then it is most likely that she was counselled for child immunization during pre-post-delivery care, which may also increase the awareness about immunization and hence coverage of vaccines. On the other hand, if the accessibility of health facility for the mother of the child is easily available and accessible, the interpersonal communication also increases. Not only that, regular communication with health care facilitators also increases the level of awareness about child immunization.

Thus, a primary objective of this study is to understand how underlying socio-economic disadvantage at the individual and household levels among social groups may influence the uptake and choice of health services. Uptake and choice (the main variables in this study) can be seen as outcome indicator coverage of full immunization within four models (Models 1-4, more information is provided in the Methodology chapter).

Hypotheses of the study

Two hypotheses ‘*characteristics hypothesis*’ and ‘*interaction hypothesis*’ have been used to elucidate the pathways and mechanisms of influence of household social affiliation (caste/tribe) on intake of child immunization in the studied states. First, the ‘*characteristic hypothesis*’ emphasizes the important role that socio-economic and demographic factors play in influencing coverage of child immunization and differentials across social groups. Second, social exclusion (differences between social groups) could be reduced through effective awareness generation and higher access to health facility among marginalized sections (*interaction hypotheses*). Detailed discussion of two hypotheses is made below.

The ‘*Characteristics hypothesis*’ states that the social group affiliation (caste/tribe) by itself has little or no independent influence on child immunization, rather it is the differences in the socio-economic and demographic composition of social groups that largely accounts for the observed differences in child immunization. If such socio-economic and demographic factors vary substantially across caste/tribe groups, coverage of child immunization level could also differ by caste/tribe. Such an impact would depend on the strength of association between socio-economic and demographic characteristics and child immunization and the degree of inter-caste/tribe variation in these characteristics. It is, therefore, necessary to examine the characteristics hypothesis and ascertain whether observed differentials in child immunization by caste/tribe are caused by variations in characteristics or are the effects of caste/tribe affiliation *per se*. In order to do so, the magnitudes of net differentials after controlling or adjusting for the effects of socio-economic and demographic characteristics must be computed. If differentials are observed even after controlling the effects of other socio-economic and demographic characteristics, one must address the *interaction hypothesis* or *assimilation hypothesis*.

The health belief model (HBM) itself calls for the interaction hypothesis as it says that there are some factors in the community that interact with the household and individual level characteristics and force a positive outcome. Therefore, the interaction between the social groups and programme factors could also affect child immunization coverage, that is, the higher the exposure to cues to action factors (programme factors, for example, mass media, institutional delivery and access to health facility), the higher immunization coverage among social groups and larger the differences in immunization coverage between SCs/STs and OCs. Or, lower the exposure to cues to action factors

(programme variables such as exposure to mass media, institutional delivery and access to health facility) among social groups, lower the immunization coverage and smaller the differences between SCs/STs and OCs. Since caste/tribe also influences immunization coverage through the provision of a platform for social interaction and diffusion of health care needs (Hummer *et al.*, 2004, Rostas, 1999), the prescription or proscription of certain lifestyles as well as the regulation of health-related behaviour of its adherents may lead to the adoption of health damaging or health promoting behaviour, thereby impacting child immunization coverage. Thus, this study seeks to examine the relationship between social affiliation and child immunization in the context of these two hypotheses in Rajasthan, Odisha and Madhya Pradesh using the data from NFHS-3. The ‘characteristics hypothesis’ emphasises whether the differentials in the child immunization coverage among SCs, STs, OBCs and OCs are attributed to socio-economic factors, and interaction hypothesis, that is, interaction between social groups and cues to action variables (programme factors) could affect child immunization coverage.

DATA AND METHODS

Sources of data and sampling

The NFHS-3, a large-scale survey conducted in 2005-06, provides the principal data for the study. It was the outcome of the collaborative efforts of many organizations such as the United States Agency for International Development (USAID), the United Kingdom Department for International Development (DFID), the Bill and Melinda Gates Foundation, UNICEF, UNFPA and the Government of India. However, technical assistance was provided by Macro International, Maryland, USA. The International Institute for Population Sciences (IIPS), Mumbai, was designated as the nodal agency for conducting, monitoring and disseminating the results of the third round of National Family Health Survey. The survey provides essential state and national level data to monitor health and family welfare programmes and policies implemented by the Union Ministry of Health and Family Welfare and other ministries and agencies.

The first round of NFHS was conducted in 1992–1993 and the second round in 1998–1999. Unlike the earlier rounds of NFHS surveys, however, NFHS-3 interviewed men aged 15-54 and never married women aged 15-49, as well as ever-married women. The NFHS-3 collected information from a nationally representative sample of 1,09,041

households, 1,24,385 women aged 15–49, and 74,369 men aged 15–54 in all 29 states covering 99 per cent of India’s population living in all the 29 states. The household response rate was 98% and the individual response rates were 94 and 87 per cent among eligible women and men, respectively (for details, see IIPS, Mumbai and ORC Macro, 2007a,b. NFHS–3: India).

Sample design

A multistage stratified sampling method was used to create a sample, representing individuals from all the 29 Indian states. The urban and rural samples within each state were drawn separately and a uniform sample design was adopted in all the states. In each state, the rural sample was selected in two stages, with the selection of Primary Sampling Units (PSUs) which are villages, with probability proportional to population size (PPS) at the first stage, followed by the random selection of households within each PSU in the second stage. In rural areas, a list of villages from the 2001 Census served as the sampling frame and it was stratified by a number of variables. The first level of stratification was geographic, with districts being subdivided into contiguous regions. Within each region, villages were further stratified using selected variables from the following list: village size, percentage of males working in the non-agricultural sector, percentage of population belonging to scheduled castes or scheduled tribes, and female literacy. In urban areas, a three-stage procedure was followed. In the first stage, wards were selected with PPS sampling. In the next stage, one census enumeration block (CEB) was randomly selected from each sample ward. In the final stage, households were randomly selected within each selected CEB. In every state, mapping and household listing operations were carried out in each sample area. The listing provided the necessary frame for selecting households at the second stage. The households to be interviewed were selected with equal probability from the household list in each area using systematic random sampling. Sample weights for households and women have been calculated to adjust for the effect of differential non-response in different geographical areas. After adjustment for non-response, the weights are normalized so that the total number of weighted cases is equal to the total number of un-weighted cases (for details see IIPS and ORC Macro.2007a,b. NFHS–3, India, 2005-06: pp. 1-12).In NFHS-3, the data were collected using three types of questionnaires: the Household Questionnaire, the Women’s Questionnaire and Men’s Questionnaire. The Questionnaires for each state were bilingual, with questions in both the principal language of the state and English.

Study population and sample size

Data used for the present study are confined to three major EAG states, viz. Madhya Pradesh, Odisha and Rajasthan. They were chosen because of their diversity in terms of social group composition, different ecological back drop and poor health status of children (infant and child mortality rates are highest in Odisha and Madhya Pradesh) and coverage of EPI vaccines is lowest to medium. In Madhya Pradesh, Odisha and Rajasthan, the NFHS-3 collected information from a representative sample of 5488, 3920, 3282 households respectively. From the sub-sample of households a total of 6427, 4540 and 3892 women aged 15-49 years; and 2725, 1592 and 1471 men aged 15-54 were interviewed to obtain information on population, health and nutrition in the state. The overall household response rates was 99 per cent in each state, and the individual response rates ranged from 98 to 99 per cent for eligible women and 93 to 98 per cent for eligible men (for details, see IIPS, Mumbai and ORC Macro, 2007c,d,e). Fieldwork was conducted from November 2005 to May 2006 (Phase 1) in Odisha and Rajasthan, and from April 2006 to August 2006 (Phase 2) in Madhya Pradesh. Like NFHS-1 and NFHS-2, NFHS-3 was designed to provide estimates of important maternal and child health indicators. In the NFHS-3 survey, information for children was collected for those who had taken birth during the five years preceding the survey, from de-facto women aged 15–49 years. In the three states, a total of 5,887 children had taken birth during five years preceding the survey.

Unit of analysis

Coverage of immunization was studied among children aged 12-23 months. However, the total number of children aged 12-23 in each state was small for a meaningful multivariate analysis. Therefore, to study the role of social group affiliation on immunization coverage after controlling for other socio-economic and demographic characteristics along with programme characteristics, the current study has focused on children *aged 12 to 59 months* (who were born during five years preceding the Survey) among de-facto women aged 15–49 years in the three major Empowered Action Groups of India, viz., Madhya Pradesh, Odisha and Rajasthan analysis.

Previous research has also examined the immunization coverage using data on living children only, due to the lack of immunization information for children who died (e.g., Pebley, Goldman, and Rodriguez, 1996). The limitation of immunization estimates

for living children has resulted in an overestimate of immunization coverage, which does not arise when using NFHS data set. Moreover, children who have not completed their first birth day were excluded from the study, because they may not have received all the recommended doses. The sample also excludes children whose mothers died as well as those whose mothers were younger than 15 years or older than 49 years at the time of survey. The cases with missing information are also excluded from this analysis. Out of a total of 5,887 children who had taken birth during five years preceding the survey in the three study states, valid data on immunization were available for 5,405 children (1,519 in Rajasthan, 1,366 in Odisha and 2520 in Madhya Pradesh).

Variables used for analysis

Several dependent variables and explanatory variables/measures have been used for different analysis in the study. Brief description and operational definition of each of dependent and explanatory variables is provided in Table 3.1. Most explanatory variables were decided a priori. In most cases, the justification for inclusion is obvious and needs no explanation. The selection of the explanatory variables (household, individual and programme levels) was based on their theoretical and empirical importance as borne out by the national and international literature on the coverage of child immunization services on the one hand, and their availability in the NFHS data set on the other. A few variables were added following close inspection of the data set, provided there was evidence for a plausible association with the outcome. Child, parental and household-level characteristics were all considered. For purposes of analysis, numerical variables were grouped into categories, and a reference group was defined.

Dependent variables

A number of dependent and independent variables have been used for different analyses and a brief description and operational definition of each one are provided in Table 3.1. In most cases, the justification for inclusion is obvious and needs no explanation. However, the rationale has been discussed in brief.

Source of vaccination data: The NFHS-3 collected information (from Q509 to Q515) on vaccination coverage for all living children born since January 2000 in Odisha and Rajasthan or since January 2001 in Madhya Pradesh). Data on vaccination are derived both from vaccination card, when the mother has it and from the mother's memory when

she cannot show it. In the survey mothers were asked whether they had a vaccination card for each child born. If a card was available, the interviewer was required to carefully note the day, month and year when each vaccination was received. For vaccinations not recorded on the card, the mother's report that the vaccination was or was not given was accepted. If she could not show a vaccination card, she was asked whether the child had received any vaccination. If it had been received, she was then asked whether the child had received vaccination against each of the six diseases. For DPT and polio, information was obtained on the number of doses of the vaccine given to the child. In such cases, mothers were not asked the dates of vaccinations. To distinguish Polio 0 (polio vaccine given at the time of birth) from Polio 1 (polio vaccine given about six weeks after birth), they were also asked whether the first polio vaccine was given just after birth or later. Responses from both the vaccination records and mothers' recall were recorded as 'yes', 'no' or 'don't know'. Goldman and Pebley (1994) demonstrate that inclusion of maternal recall data improves the accuracy of estimates of immunization coverage, although they are subject to recall error. Therefore, both mother's recalls and card coverage data are compiled. Around four per cent of data observed non-numeric, i.e., mothers responded 'Do not know' (DK) for the specific vaccines and which cannot be distributed either as card coverage or mother's recalls coverage. Hence, children in these cases are considered as not having received the vaccine.

Immunization status: To understand the coverage of full immunization a composite variable was prepared taking into consideration eight doses of specific vaccines, one dose each of BCG and Measles and three doses each of DPT and Polio as per WHO guidelines. Immunization status was categorised into fully immunized, if a child at any time before his/her first birth day received all the eight doses, otherwise termed as not fully immunized, in case of analysis of immunization status (dependent variable) with two category variables.

Immunization status is categorised into three categories as, (a) if a child at any time before his/her first birth day received all the eight doses, he was are considered as fully immunized, (b) if a child who missed at least one dose out of the eight vaccines was considered as partially vaccinated; and (c) if a child who did not receive any dose of the eight vaccines was considered as unvaccinated/not vaccinated, in the case of analysis of immunization status with three categories (dependent variable).

Vitamin A Supplements (VAS): The NFHS-3 collected information on the consumption of vitamin A-rich foods and on the administration of VAS for the youngest child aged 6-59 months living with the mother. It was derived both from vaccination cards, when the mother had one, and also from the mother's memory when she could not show it. If a card was available, the interviewer copied the date for doses of VAS. Then, the mother was asked whether the child had received VAS. VAS is categorised into child received VAS (both from card and mothers recall) as one, otherwise zero.

Analysis of coverage of full immunization and VAS coverage are restricted to living children who had reached their first birth on the date of interview due to three reasons. First, children under age one are not yet eligible to receive all the doses of vaccines; second, the analysis focuses on vaccination status on or before first birth day, not on age appropriate vaccination, and finally, the information on vaccination coverage for children who died was not obtained from data base.

Explanatory variables

Social group affiliation is the central variable of interest in the analyses. In addition to the social group affiliation (caste/tribe), following other modifying variables (household, parental, child level characteristics) and cues to action variables (programme factors) were all considered. The modifying variables include household characteristics (i.e., residence, household structure, household's standard of living index); characteristics of parents (father's education, mother's education¹, mother's work status, mother's age at the time of childbirth) and characteristics of children (age, birth order and sex) were used as explanatory/independent variables in the three states. Besides modifying variables (socio-economic and demographic characteristics at household, parental and individual),

¹Although women's level of education and husbands' education are related, inclusion of both variables in the multivariate analysis does not lead to multi-collinearity as all other effects do not alter considerably. Similarly, in the case of mothers' age at child birth and parity also. Therefore, both variables are included in the analysis and show a significant effect. However, multivariate results show that the effect of the first variable is stronger. Potential problems of multi-collinearity acknowledged and assessed the multi-collinearity between the variables [values from Pearson correlation coefficients(r) examined] and highly correlated variables were eliminated from the models.

three proxy programme exposure variables (i.e., exposure to electronic media, place of delivery and access of health facility to mother) which influence immunization uptake were used as cues to action variables, from the available NFHS data (Table 3.1). Since the NFHS data do not provide direct indicator for perceptions, the outcome of perception may be determined from changes in outcome. Further, using the available data on women's constrains/barriers to access health facility in the NFHS-3, a composite index on mother's access to health facility (with three categories low, medium and high access) was constructed with the Principal Component Analysis.

Most of the explanatory variables such as caste/tribe, household structure, standard of living (index), parents education (mother's and father's), mother's work status, current age of the child, birth order and sex) are used without distorting the original nature as given in the NFHS-3 data files (Table 3.1). However, for a meaningful analysis of data, some of the variables (cues to action variables or programme exposure factors) have been generated or recoded with the relevant data by keeping unchanged the original one from the NFHS-3 data files. Age of mother at the time of birth of index child and three cues to action variables such as exposure to electronic media, place of delivery and mother's access to health facility) included in the analysis, since they have marked influence on immunization uptake.

In the following paragraphs, a short description on data source of generated variables (exposure to electronic media, place of delivery and mother's access to health facility), standard of living index and other major variables of interest and social group affiliation are provided.

Social group affiliation (Caste/tribe): Social group affiliation is the principal variable of interest in the analysis. The NFHS-3 data collected information about social affiliation of the household head (Q45), women (Q117), and men (Q119). If the response to these questions was either caste or tribe, the name of the caste or tribe was written down by the interviewer (Q46, Q118, and Q119 of the household, woman and man's questionnaires respectively). From this information, one of the four castes categories such as, SC, ST, OBC and those who are neither SC nor ST nor OBC and are designated as OC, are considered for the analysis. Depending on the nature of analysis, the social group affiliation variable categorized with 2 to 4 different combination of categories (specifications furnished in Table 3.3). Vast majority of the children aged 12-59 months

in the study states (97, 91 and 88 percent in Odisha, Madhya Pradesh and Rajasthan) belong to Hindu religion therefore, religion variable not focused in the analysis and extensively assessed the role of social group affiliation on childhood immunization coverage in the study states.

Household standard of living (Index): In the NFHS–3, information on household income or expenditure was not collected. However, detailed data on housing conditions and ownership of certain assets were obtained and based on the information, a summary household measure called Standard of living index (SLI), is calculated and available in the data files. The sum of all weights for each individual (woman) indicates her position in standard of living in terms of possession of household durable goods and household amenities. Based on these total scores, the households were classified into three categories as having ‘low, medium, and high’ standard of living (IIPS and Macro International, 2007).

Mother’s age at birth of index child: This variable is computed based on the difference between the date of birth of the index child and woman’s date of birth (B3 and V011 of NFHS–3 data files). Analysis of age of mother at birth of index child, i.e., on the timing of birth of index child, is carried out to see whether the attainment of motherhood is closely linked to child immunization. It may affect the frequency of complications during pregnancy and childbirth and because it may be related to the woman's attitudes about health care providers.

Exposure to media: In the NFHS following four questions were asked to mothers to assess their exposure to electronic and print media. (a) “Do you read a newspaper or magazine almost every day, at least once a week, less than once a week or not at all?”; (b) “Do you listen to the radio almost every day, at least once a week, less than once a week or not at all?”; (c) “Do you watch television almost every day, at least once a week, less than once a week or not at all?”; (d) “Do you usually go to a cinema hall or theatre to see a movie at least once a month?” For the analysis, responses from the questions (b) and (c) on exposure to electronic media are considered and ‘those who ever watch television or listen radio were considered as exposed to electronic media.

Place of delivery: NFHS had asked women who had delivered during the preceding five years of the survey, what was the place of delivery. Responses include home delivery,

delivery in government institute, private institute, NGO and any other. This information recorded and categorized into two categories, as delivery at home and delivery at any institutional facility and used for the analyses.

Women's access to health facility (index): Access to health care is a complex multi-dimensional concept and it reflects the coverage of health facilities also. The NFHS-3 asked questions (Q557a-h) to all women about eight situations that could prevent women to access medical advice or treatment for themselves and coded as either a big problem or a small problem or no problem. Questions covered (Q557a-h) are read as follows.

Now I would like to ask you some questions about medical care for yourself.

Many different factors can prevent women from getting medical advice or treatment for themselves. When you are sick and want to get medical advice or treatment, is each of the following a big problem, a small problem, or no problem?

(a) Getting permission to go; (b) Getting money needed for treatment; (c) The distance of health facility; (d) Having to take transport; (e) Finding someone to go; (f) Concerned no female health providers available; (g) Concerned no health providers, and (h). Concerned that no drugs available.

To assess the perceived constraints/barriers of women in accessing health care, the responses of these eight questions on access to health facility are used for computation of 'access score' with Principal Component Analysis. Analysis conducted using all women sample and data are consequently weighted with women weights. A higher score on this denotes higher access to health facility (no/least constraints/barriers) and a lower score implies least (high constraints/barriers) to access health facility. The composite access score used to construct/recode as 'women's access to health facility index' and categorized into three categories as 'low, medium and high'. Detailed description on justification, steps involved in the construction of the index, and results are discussed in Section 2 of Chapter 4. The set of eight responses used in quantifying the access to health facility is shown in Table 3.2. However, the questions did not specify any particular source of care but pertained to wherever the woman would seek care.

Analysis of data

Methods of analysis

The conceptual framework calls for examining the role of social group affiliation (caste/tribe) after controlling the other socio-economic and demographic characteristics of mothers and children along with and programme factors on child immunization coverage in the three states. For this purpose, both bi-variate and multivariate analysis have been carried out. The association between outcome and exposure variables was assessed using cross-tabulations with Pearson's chi-square (χ^2) analysis, Spearman's correlation for ranked variables along with reliability tests and significance for all analysis was set at $p < 0.05$. Only variables show a significance level of $p < 0.10$ in the bivariate analysis considered for inclusion into the multivariate analysis. Independent variables for multivariate analysis were chosen based on prior findings and the variables that did not contribute significantly to the outcome were dropped from the model. Related variables were also checked for correlation to avoid problems of multi-collinearity and unstable regression models. Detailed information on methods adopted for analysis of data is presented in Table 3.3. First, the gross differentials in source of vaccination data, coverage of specific vaccination and immunization status, drop-out rates during multiple doses of vaccines, missed-out of only one dose of vaccination, coverage of DPT-3, measles, vaccination/immunization status, coverage of vitamin A supplements across the four caste/tribe groups are examined through bivariate analysis.

Later, in order to examine the net effects of social group affiliation (caste/tribe) on the dependent variables (DPT3, measles vaccination and immunization status), controlling for all other modifying and cues to action factors, different multivariate techniques (logistic and multinomial logistic regression analysis) were used. Dependent variables which have dichotomy in nature are analysed with binary logistic regression, whereas dependent variables, which have more than two categories are analysed with multinomial logistic regression. Demographic, socio-economic characteristics and programme variables that were statistically significant in the bivariate analysis ($p < 0.05$) were entered into four regression models (Model 1-4) for all analysis, in line with the HBM conceptual framework, as mentioned in the Table 3.3. Briefly, Model 1 assesses the effect of caste/tribe only; Model 2 contains caste/tribe (Model 1) along with three

programme variables (cues to action); Model 3 caste/tribe (Model 1) along with other modifying variables (socio-demographic variables); and Model 4 contains all the variables from the three previous models (Model 1 + Model 2 + Model 3). In the similar lines, the effects of coverage of full immunization and social determinants, particularly the role of social group affiliation (caste/tribe) on Vitamin A supplementation coverage are examined in four models, sequentially controlling for other risk factors. However, full immunization (Model 1), and social group affiliation (Model 2) along with social only determinants (Model 3) considered.

Finally, to examine whether the social group affiliation interacts with programme characteristics, i.e., the interaction effects of caste/tribe with and three programme variables by creating new interaction variables (viz., ‘Caste/tribe and exposure to electronic media’, ‘Caste/tribe and place of delivery’ and ‘caste/tribe and access to health facility’) on child immunization were assessed. It, it is possible that at least a portion of differentials in the child fully immunized among social groups are attributable to differences in the programme related characteristics. The effect of these new variables on child vaccination is accessed using multivariate logistic regression analysis to quantify the net effect of ‘Caste/tribe and exposure to media’, ‘Caste/tribe and place of delivery’ and ‘caste/tribe and access to health facility’, independent of other background factors on the full immunization. It should be noted that since caste/tribe and programme variables are categorized, inclusion of two interaction variables in a single analysis brings in multicollinearity. For example, if ‘Caste/tribe and exposure to electronic media’, ‘Caste/tribe and place of delivery’ and ‘Caste/tribe and access to health facility’ are included in a single analysis, the matrix becomes singular and hence, both cannot be examined in one analysis. Therefore, three separate analysis are carried out for three interaction variables controlling other socio-economic variables and relevant programme variables with four models (Models 1-4) in line with the HBM conceptual framework (more detailed information on data and methods discussed in the Chapter 7). Data analysis was conducted using SPSS software version 21.0 (IBM SPSS statistics standard, USA) and STATA 11.2 (StataCorp, USA) software.

As the emphasis is primarily on the role of social group affiliation (caste/tribe), results are discussed in detail on the net effect of caste/tribe and other socio-economic, demographic and programme factors on the dependent variables. Other explanatory

variables have been used only to substantiate the findings or to develop the model, and are not discussed in detail, except for major social determinants, which are strongly significant and most important predictors on the model.

Brief description of statistical techniques used for the analysis

The analysis has employed different bivariate and multivariate statistical techniques. Many of the techniques are well known. Nevertheless, details of the statistical techniques used for the analysis of data in following chapters, the need to use the specific techniques, and basic model has been briefly provided in the following lines. However, formulae and algorithms are not described. Appropriate references have been cited for these.

Drop-out rates/missed-out rates: definitions, rationale and calculations

In order to achieve universal immunization, it is important to track children so that all vaccines can be administered. It is also important to understand at what stage children drop-out from receiving all the required vaccines. A dropout is defined as a child who failed to return for subsequent doses for which he or she is eligible (Eduard Bos and Amie Batson, 2000). A high drop-out rate means that either quality of immunization services is very poor or mothers have poor access to them. In order to see the magnitude of drop out between vaccinations, drop-out rates are calculated as the difference in percentage coverage in between the consecutive vaccines divided by coverage of the first dose of vaccination.

Similarly, DPT and polio dropout rates between the successive are defined and calculated. For instance, drop-out rates

Between DPT1 and DPT2: $(DPT1 - DPT2) * 100 / DPT1$

Between DPT2 and DPT3: $(DPT2 - DPT3) * 100 / DPT2$

Between DPT1 and DPT3: $(DPT1 - DPT3) * 100 / DPT3$

Principal component analysis (PCA)

PCA is a multivariate statistical technique used to reduce the number of variables in a data set into a smaller number of ‘dimensions’ (Vyas and Kumaranayake, 2006; Rutstein and Johnson, 2004). In mathematical terms, from an initial set of n correlated variables, PCA creates uncorrelated indices or components, where each component is a linear weighted combination of the initial variables. For example, from a set of variables X_1 through to X_n ,

$$\begin{aligned} PC_1 &= a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n \\ &\vdots \\ PC_m &= a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n \end{aligned}$$

Where a_{mn} represents the weight for the m^{th} principal component and the n^{th} variable.

The weights for each principal component are given by the eigenvectors of the correlation matrix, or if the original data were standardized, the co-variance matrix. The variance (λ) for each principal component is given by the eigenvalue of the corresponding eigenvector. The components are ordered so that the first component (PC1) explains the largest possible amount of variation in the original data, subject to the constraint that the sum of the squared weights ($a^2_{11} = a^2_{12} + a^2_{12} + \dots + a^2_{1n}$) is equal to one. As the sum of the eigenvalues equals the number of variables in the initial data set, the proportion of the total variation in the original data set accounted by each principal component is given by λ_i/n . The second component (PC2) is completely uncorrelated with the first component, and explains additional but less variation than the first component, subject to the same constraint. Subsequent components are uncorrelated with previous components. Therefore, each component captures an additional dimension in the data, while explaining smaller and smaller proportions of the variation of the original variables. The higher the degree of correlation among the original variables in the data, the fewer components required to capture common information.

In the study, Principal component analysis (PCA) is used to construct composite index on women’s access to health facility. Detailed description on the steps in constructing access to health facility index viz., selection of access variables, application

of PCA, interpretation of results and classification of women according to level of access to health facility (low, medium or high) are provided in Chapter 4, Section 2.

Logistic regression analysis

Logistic regression (logit regression) is used when the response or dependent variable is dichotomous (i.e., binary, or 0-1). The predictor variables may be quantitative, categorical or a mixture of the two. Suppose, the probability of the occurrence of event Y, $[P(Y=1)]$ depends on a set of explanatory variables $X_1, X_2, X_3, \dots, X_k$.

The basic form of the logistic function is

$$P = \frac{e^Z}{1 + e^Z}$$

where Z, is a linear function of a set of predictor variables, $X_1, X_2, X_3, \dots, X_k$, given by

$$Z = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$$

p is the estimated probability of having a vaccination card or receiving DPT 3, or measles or full vaccination.

Logit of P is derived by taking natural logarithm, that is, $\log \left[\frac{p}{1-p} \right] = Z$.

The quantity $\left[\frac{p}{1-p} \right]$ is called the odds and hence $\log \left[\frac{p}{1-p} \right]$ the log odds (For details, see Kendall 1975; Fox 1984). The coefficients $b_0, b_1, b_2, \dots, b_k$ are similar to regression coefficients and are called logit regression coefficients.

Owing to the dichotomous nature of the dependent variables viz., availability of vaccination card (yes, no), received specific vaccinations –DPT3 and measles (yes, no), received VAS (yes, no), the technique of multiple logistic regression analysis has been adopted for the analysis. The independent variables are recoded into categorical indicator variables. One value of each variable is chosen to be the reference category. The reference category was the first category of the variable. For ease of interpretation, the magnitude and direction of association between the variables were expressed as adjusted odds ratios (OR). The latter is a measure that approximates how much more likely, (or

unlikely) it is for the outcome, in this case, for example, being vaccinated with DPT3 or measles vaccination, to be present among those with a given attribute relative to the reference category. The odds ratio for the reference category is equal to 1.0. If an odds ratio is greater than 1.0 this indicates an increased likelihood of the event occurring, while an odds ratio less than 1.0 indicates a decreased likelihood of its occurring. An odds ratio of 1 means that, the variable has no effect. For continuous variables the OR measures the change in the dependent variable per unit change in the variable. A variable was considered significantly associated with mortality when its *p* value was below 0.5.

Multinomial logistic regression analysis

Multinomial logistic regression analysis used to test for significant associations with women’s access to health facility (low, medium and high access) and coverage of immunization status (Full, partial and not immunized) controlling for other covariates, with high access to health facility and full immunization as the reference category.

The multinomial logistic regression model is an extension of the binary logistic regression model where the dependent variable is polytomous, i.e., its values consist of more than two categories (Hosmer and Lemeshow, 2000; Long and Freese, 2001). The basic assumption of multinomial logistic regression model that should be strictly fulfilled is that the categories of the response variable should be mutually exclusive and exhaustive, i.e., a sample member must fall in one and only one of the categories. The above assumption is fulfilled in the analysis. In this model the response variable is mutually exclusive and exhaustive.

The Multinomial Logistic model can be given as:

$$\log \frac{p_1}{p_2} = a_1 + \sum b_{1i} x_i, i= 1, 2, 3, \dots, n;$$

$$\log \frac{p_2}{p_3} = a_2 + \sum b_{2i} x_i, i= 1, 2, 3, \dots, n;$$

and

$p_1 + p_2 + p_3 = 1$. Where a_1 and a_2 are constants and b_{1i} , b_{2i} are the coefficients of x_i 's.

For example in the present analysis, p_1 is the estimated probability of low level of women’s access to health facility/not vaccinated; p_2 denotes the estimated probability of medium level of women’s access to health facility/partially immunized or vaccinated; and

p_3 is the probability of high level of women's access to health facility/ received full immunization or vaccination. Here p_3 is the reference category. For interpretation of results the relative risk ratios (RRRs) which is also an OR, but is instead placed in a risk discourse are used.

The RRRs is an OR, but for the specific comparison of the outcome in question against the chosen base case. The OR of a coefficient indicates how the risk of the outcome falling in the comparison group compared to the risk of the outcome falling in the referent group changes with the variable in question. An $OR > 1$ indicates that the risk of the outcome falling in the comparison group relative to the risk of the outcome falling in the referent group increases as the variable increases. In other words, the comparison outcome is more likely. An $OR < 1$ indicates that the risk of the outcome falling in the comparison group relative to the risk of the outcome falling in the referent group decreases as the variable increases. In general, if the $OR < 1$, the outcome is more likely to be in the referent group.

Quality of data

The NFHS-3 used a uniform sample design, questionnaires, field procedures to ensure the highest possible data quality. Special attention was also paid to missing information, skip instructions, filter questions, age information, completeness of the birth history and the health sections.

In any Survey or Census, ages are poorly reported, particularly heaping on ages 0 and 5. However, the NFHS-3 age data are of considerably better quality than other sources. The extent of missing information is very low for background characteristics, and demographic and health questions/measures. For instance, missing information is only 1.23 per cent for the month of birth of children born and 0.17 for the month and year (out of a total 180641) in the past 15 years in the total sample (IIPS and ORC Macro, 2008b:34). Similarly, for data on education, missing information is only 0.01 and 0.02 per cent for women age 15-49 (out of a total of 124,385 cases) and men age 15-54 (out of a total of 74,369), respectively.

Despite these measures to improve data quality, NFHS-3 is subject to the some errors that are inherent in all retrospective sample surveys—namely, the omission of some births (especially births of children who died at a very young age) and the difficulty

of determining the date of birth of each child accurately. These difficulties are likely to somewhat bias the results.

Limitations of the data

This study has several strong points in its large sample size and quality assurance of data management. However, there are few important limitations which may bias the findings. Following five limitations have to be taken into account when interpreting the results of this study. Among the major limitations first, the study is solely based on NFHS-3 data and the data represents a decade old (conducted during 2005-06), i.e., the immunization data of the study refers to the children who have born five years preceding the survey. Although the National Rural Health Mission (NRHM) activities were launched in the study states in 2005 and the impact of NRHM interventions should have been reflected through increase in the coverage of immunization. However, the increase in the coverage of immunization is invisible in these 14 years, for instance, as compared to the NFHS-3 immunization coverage rates, findings from the other large scale surveys conducted during the period of 2007-2012 Viz., DLHS-RCH-3 (2007-08), CES (2009, 2011), AHS (2011-12) reports lower coverage of immunization in the three study states/districts and particularly for social groups. Findings from DLHS-RCH, CES and latest AHS, shows that the increase in the coverage of full immunization (MP and Odisha at state levels and among SC/STs in the state of Rajasthan) is not more than 10 percent during the last 10-14 years (Tables 1.2 and 1.3). Therefore, due to the non-availability of appropriate and reliable data on immunization at state levels and the subsequent survey, the NFHS-4 is still under implementation (2014-15) in the study states and India, and thus unit level data are yet to become available, still the NFHS-3 data provides most reliable source of information and gives a good opportunity to examine extensively the social determinants of immunization coverage in the study states with application of advanced statistical analytical techniques with specific analyses. Therefore, for understanding of the study objectives and to achieve objectives of the study, we believe that the analyses based on the NFHS-3 still be useful programmatically.

Secondly, since the study is based on NFHS-3, the analytical framework designed within the NFHS-3 framework and utilized the available data collected in the survey and not possible to go beyond to accesses the deterministic and associational variables etc. relating to immunisation. Specifically, the vaccination status of children did not take into

account valid or invalid doses (e.g. doses beyond the appropriate vaccination age for a given dose). *Also*, the vaccination status was ascertained as documented in the health cards of the children and by caregiver's recall, which may introduce some bias (over-estimating vaccination in some cases and under-estimating vaccination in others), because mother may forgot the total doses of vaccine that the child took. *Third*, children who have died did not have vaccination status data in the datasets. *Fourth*, the women's access to health facility index is generally only recommended for use as a ranking mechanism, not as an absolute measure. It is limited in its ability to measure multiple dimensions of women's access to health facility (getting medical advice or treatment for themselves), and levels in those dimensions. *Finally*, a large number of variables and models have been used which may increase the number of occasions where, in spite of no true underlying relations observed, or became statistically significant. Despite the above limitations, the findings are important to understand factors associated with full immunization coverage among social groups.

Table3.1 List and description on dependent and independent variables used for the analysis

Dependent variables	Definition and description	Categories
Availability of vaccination card	Availability of vaccination card for immunization data	Yes, No
Coverage of specific vaccinations	Received eight doses of specific vaccinations each of BCG, measles, and three doses each of DPT and polio vaccine	Yes, No
Vitamin A supplementation coverage		
Drop-out rates between vaccinations	A drop out is defined as a child who failed to return for subsequent doses for which he or she is eligible	Drop-out rates between DPT1-3 and Polio 1-3 vaccinations
Immunization coverage status		<i>Full:</i> received all eight doses of vaccinations; <i>Partial:</i> missed at least one dose of the eight vaccine; <i>No:</i> does not receive any dose of the eight vaccines
Independent variables		
Caste/tribe	Social group affiliation	Scheduled caste (SC), Scheduled tribe (ST), Other backward class (OBC), Other caste (OC)
Residence	Current place of residence	Urban, Rural
Household structure		Joint family, Nuclear family
Household standard of living (Index)	Proxy for the economic status of the household	Low, Medium, High
Mother's education	Level of education of the mother	Illiterate, Primary, Secondary +
Father's education	Level of education of the father	Illiterate, Primary, Secondary +
Mother's work status	Usual activity of the mother	Working , Not-working
Mother's age at the time of (index) child birth (in years)	Difference between the date of birth of the index child and mother's date of birth	Categorized (<20, 21-25, 26+) for bivariate analyses, and Continuous for multivariate analyses
Age of child (in months)	Current age of the child	12-23, 24-35, 36-47, 48-59
Child's birth order	Order of birth of the (index) child	1, 2, 3+
Child's sex	Sex of the (index) child	Male, Female
Mother's exposure to media	Exposure to electronic media (Radio and television)	Yes – exposed; No – Not exposed
Place of delivery	Place of delivery	Home, Institution
Accessibility health facility (Index)	Proxy for the mother's access to health facility (a composite index constructed with constrains to access health facility)	Low, Medium, High
Caste/tribe and exposure to electronic media	Interaction between Caste/tribe and exposure to electronic media	SC/ST - exposed; SC/ST - not exposed; Non-SC/ST - exposed; Non-SC/ST - not exposed
Caste/tribe and place of delivery	Interaction between Caste/tribe and place of delivery	SC/ST-Institutional delivery; SC/ST-home; Non-SC/ST-Institutional; Non-SC/ST-home delivery
Caste/tribe and access to health facility	Interaction between Caste/tribe and mothers access to health facility within three different (low/medium/high) levels	SC/ST-low; SC/ST-medium; SC/ST-high; Non-SC/ST-low; Non-SC/ST-medium; and Non-SC/ST-high

Table 3.2. Variables used for construction of composite index on women's access to health facility

Specific constrains in accessing medical advice or treatment	Codes	Categories and values used for analysis
a. Getting permission to go	PERMISSION	1=Big problem 2=Not a big problem 3=No problem
b. Getting money needed for treatment	GETTING MONEY	1=Big problem 2=Not a big problem 3=No problem
c. The distance to the health facility	DISTANCE	1=Big problem 2=Not a big problem 3=No problem
d. Having to take transport	TAKING TRANSPORT	1=Big problem 2=Not a big problem 3=No problem
e. Finding someone to go (with you)	FINDING SOMEONE	1=Big problem 2=Not a big problem 3=No problem
f. Concern that there may not be a female health provider	NO FEMALE PROVIDER	1=Big problem 2=Not a big problem 3=No problem
g. Concern that there may not be any health provider	NO PROVIDER	1=Big problem 2=Not a big problem 3=No problem
h. Concern that there may be no drugs available	NO DRUGS	1=Big problem 2=Not a big problem 3=No problem

Source: IIPS & ORC Macro (2007b:105)

Table 3.3 Methods and models adopted for the analysis of data on immunization among children aged 12-59 months in three selected states (Rajasthan, Odisha and Madhya Pradesh) of India, using third National Family Health Survey (NFHS-3)

Analysis of data: Methods and models		Variables used			
Model-1:	Caste/tribe	Effect of 4 categories of caste/tribe (SC, ST, OBC and OC) Effect of 3 categories of caste/tribe (SC/ST, OBC and OC) Effect of 2 categories of caste/tribe (SC/ST and Others)			
Model-2:	Model 1 + Program variables (Cues to action)	Caste, access to health facility (index), place of delivery, exposure to media			
Model-3:	Model 1 + Modifying variables (Other socio-demographic characteristics)	Caste, residence, household structure, household standard of living(index), father's education, mother's education, mother's work status, mother's age at the time of child's birth, birth order, and sex of the child			
Model-4:	Model 1 + Model 2 + Model 3	Caste, residence, household structure, household standard of living, father's education, mother's education, mother's work status, mother's age at child's birth, birth order of the child, sex of the child, mother's access to health facility, exposure to media, and place of delivery of the index child			
Dependant variables	Categories	Models/variables used	Caste/tribe	Techniques employed	Chapter
Women's access to health facility (Composite index) [Unit of analysis: Women aged 15-49]	Low, medium and high access to health facility (Reference: High access) No Vs. High access; Medium Vs. High access	Variables on women's constraints to access health facility and social determinants of access	4 categories	Principal component analysis, Reliability analysis, and Multinomial logistic regression	4
Vaccination card Coverage of vaccines [Unit of analysis: children age 15-59 months]	Yes/No % specific vaccinations % drop-out vaccinations % missed-out vaccination	Social determinants Caste/tribe and source of vaccination data	4 categories 4 categories	Logistic regression analysis Frequency distribution, cross tabulations, calculation of rates	5
Specific vaccinations	Received DPT-3: Yes/No Received Measles: Yes/No	Model 1, 2, 3 & 4 Model 1, 2, 3 & 4	4 categories	Logistic regression analysis	
Vitamin A Supplementation Immunization coverage status	Received - Yes/No No Vs. full immunization; Partial Vs. full immunization (Reference: Full immunization)	Model 5, 2, 3 & 4 Model 1, 2, 3 & 4	3 categories 3 categories	Logistic regression analysis Multinomial logistic regression analysis	6
Interaction effects of caste/tribe with program variables on immunization	Caste/tribe by exposure to media Caste/tribe by place of delivery Caste/tribe by access to health facility	Model 1, 2, 3 & 4 Model 1, 2, 3 & 4 Model 1, 2, 3 & 4	2 categories	Logistic regression analysis	7

Notes: Scheduled castes (SC) and scheduled tribes (ST) are identified by the Government of India as socially and economically backward and needing protection from social injustice and exploitation. Other backward class is a diverse collection of intermediate castes that were considered low in the traditional caste hierarchy but are clearly above SC. 'Other caste' is thus a default residual group that enjoys higher status in the caste hierarchy.