Better Data for More Effective Policies and Programmes

SMALL AREA ESTIMATION

Background

UNFPA aims to strengthen the capacity of countries to produce and disseminate quality disaggregated data on population and development indicators that allow for mapping of demographic disparities and social and economic inequalities, and support programming in humanitarian settings.

However, for many important indicators, such data are not available at high enough resolution. Small Area Estimation (SAE) offers a means of estimating indicators when critical data are not available at lower levels of geography (i.e. district or municipality levels).

This brochure, describes a method of SAE that combines census and household survey data. It also presents a case study based on data from Nepal. The case study illustrates use of census as well as demographic and health survey data to generate estimates of three family planning related indicators for Village Development Committees (VDCs) and Municipalities in rural and urban areas, respectively.

Advantages of SAE

- Better use of existing data within national statistical systems to identify smaller geographic areas where the needs for development are greatest
- Improved targeting of interventions and allocation of resources

How can I learn more about SAE?

Contact Rachel Snow (rsnow@unfpa.org), Chief, Population and Development Branch, Technical Division, UNFPA, if you are interested in implementing SAE in your country.

Acknowledgement: We would like to acknowledge UNFPA Nepal Country Office and the Nepal Central Bureau of Statistics for their support to the Nepal case study.

Photo credits: Cover: © UNFPA/S. Zhang; Page 4: © UNFPA/LACRO/C. Wint

January, 2020

United Nations Population Fund
605 Third Avenue, New York NY 10158
www.unfpa.org

Methodology

MAIN STEPS OF SMALL AREA ESTIMATION

Several SAE methods have been developed. This brochure describes a method that was developed by Elbers, Lanjouw, and Lanjouw, in 2003 (hereafter called the ELL method).

The ELL method is anchored on correlational analysis. It takes advantage of the correlations between variables that are commonly measured in both censuses and surveys (such as age, place of residence, education level, socio-economic status, etc.) and specific indicators of interest that are exclusively measured in surveys.

1

Census and survey data assessment and harmonization

Census and survey data assessment and processing involves identifying "common variables" associated with the key Indicator X that are contained in both survey and census data; the analysis and comparison of the distribution of each of the variables across the two datasets; and harmonization of the datasets.

2

Develop and identify the best model based on survey data

Using the survey data, the ELL method first builds a regression model that predicts the likelihood of occurrence of the event of interest among individuals in the survey, using the variables that are common between census and survey data as predictor variables. The best model is selected based on predictive power and accuracy assessment.

3

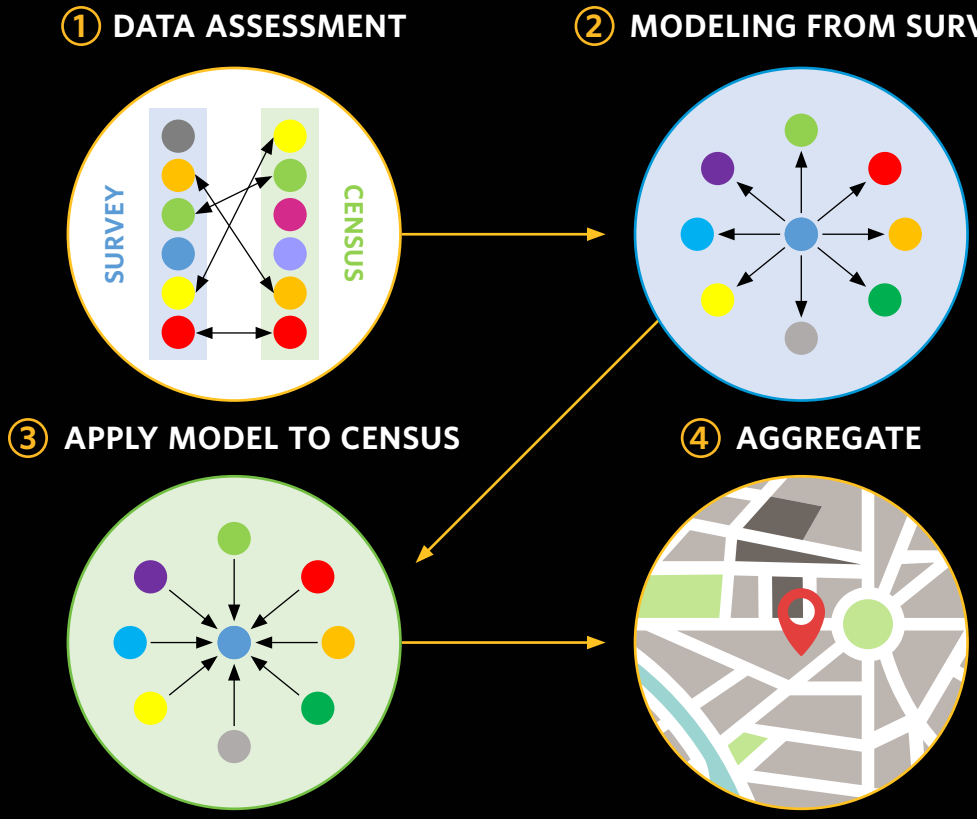
Apply the model to census data to predict individual-level estimates based on census data

The coefficients from the survey-based regression model are then applied to the "common variables" in census data to build a corresponding model to estimate Indicator X at individual level using census data.

4

Aggregate the estimates from the individual-level to any level of geography (including small areas)

The individual-level estimates can be aggregated to any unit of analysis including different levels of geography (e.g. district or municipality).


① DATA ASSESSMENT
② MODELING FROM SURVEY
SURVEY
CENSUS
③ APPLY MODEL TO CENSUS
④ AGGREGATE

Small Area Estimation (SAE) for Better Family Planning Programming

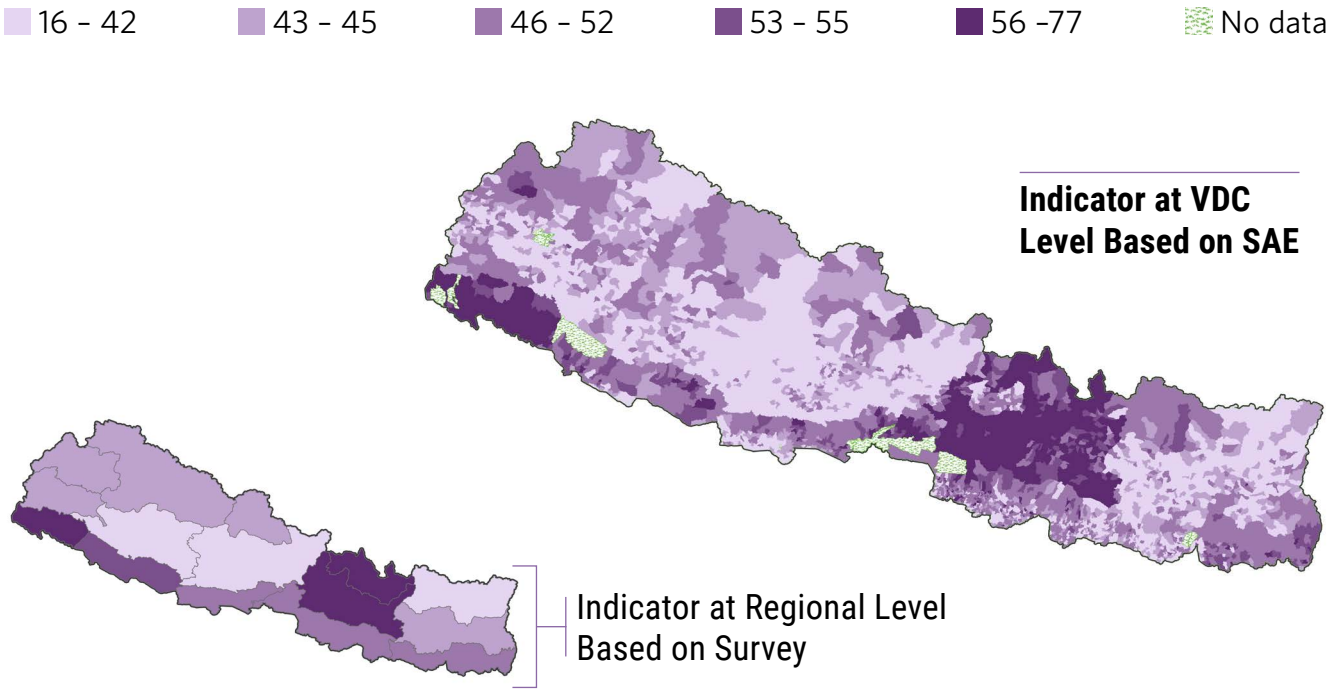
A CASE STUDY OF NEPAL

Introduction

This case study applies the ELL method to produce municipality/VDC level estimates of family planning indicators in Nepal using data from the 2011 Demographic and Health Surveys and the 2011 Population Census Public Use Microdata Sample (PUMS) data.

Individual level probabilities of contraceptive use are estimated using logistic regression models with variables that are common in the two datasets included as predictor variables.

Contraceptive prevalence rate (any method)

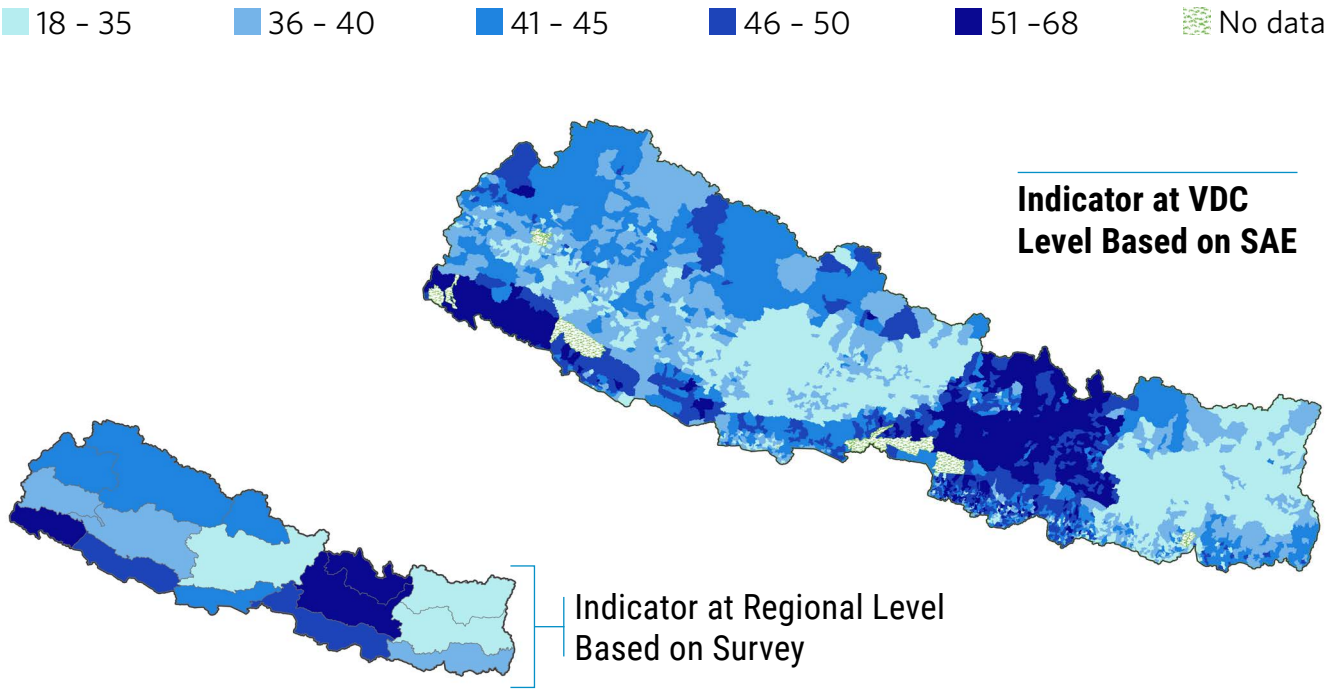


Contraceptive Use Indicators of Interest

- Contraceptive prevalence rate (CPR) — percentage of women of reproductive age (15–49 years) who are currently using, or whose sexual partner is currently using, at least one method of contraception, regardless of the method used
- Modern contraceptive prevalence rate (CPRm) — contraceptive prevalence rate due to modern methods
- Unmet need for family planning — percentage of women of reproductive age, who are sexually active and want to stop or delay childbearing, but are not using any method of contraception
- Proportion of demand for modern contraception satisfied (PDSm) / SDG Indicator 3.7.1 — percentage of women of reproductive age who have their need for family planning satisfied with modern methods

$$PDSm = \frac{CPRm}{CPR + Unmet\ need} X100$$

Modern contraceptive prevalence rate



Variable Selection

- Several variables associated with women's use of contraception can be found in census, such as age, number of children ever born, place of residence, highest level of education, marital status, relationship to household head, religion, characteristics of housing structure, household wealth and services, etc.
- The selection of variables used to predict contraceptive use is based on literature review, common variables existing in both census and survey data, and recommendations of national experts.

Logistic Regression Model for SAE

$$P = \frac{exp(\beta_0 + \sum_{i=1}^{n} \beta_i x_i)}{1 + exp(\beta_0 + \sum_{i=1}^{n} \beta_i x_i)}$$

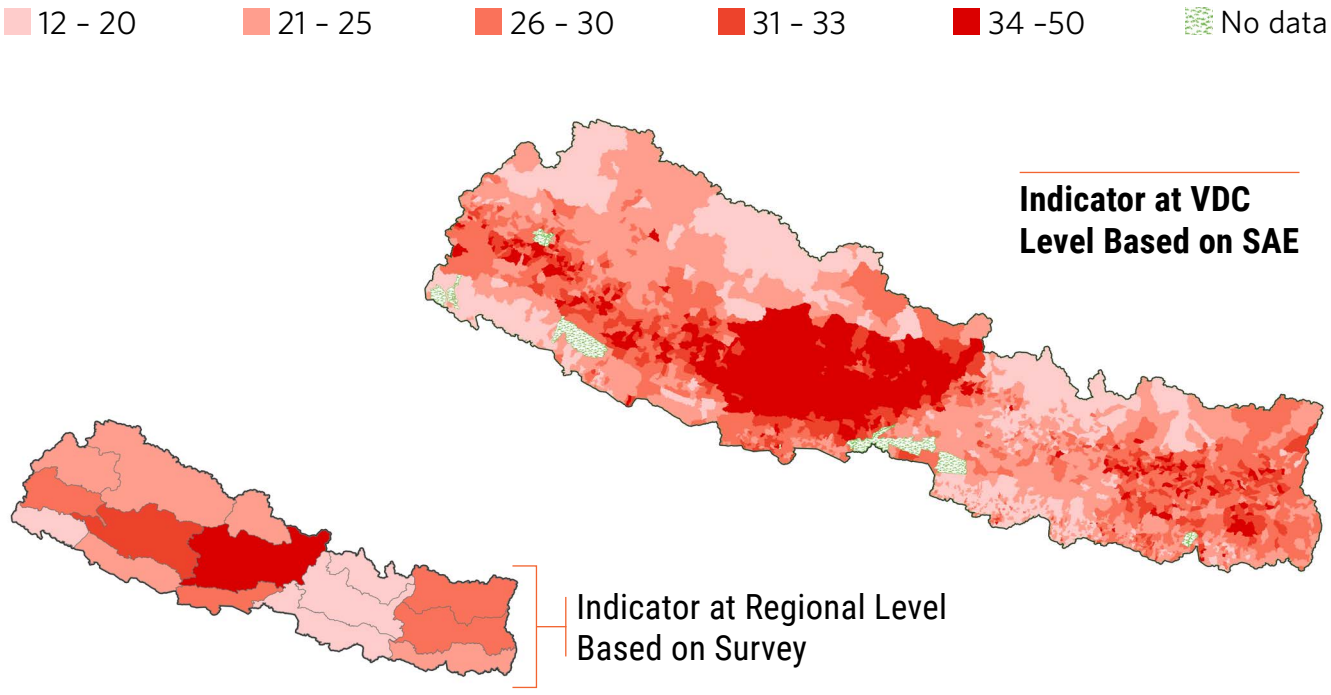
- P is the probability that an individual uses modern contraception
- $x_1, \ldots. x_n$ are the predictor variables included in the model
- Coefficients $\beta_1, \ldots, \beta_n$ are the effects for each of the predictor variables

References

Elbers, C., Lanjouw, J., and Lanjouw, P., (2003), Micro-level Estimation of Poverty and Inequality, *Econometrica, Vol.* 71, pp. 355-364.

World Bank. (2013). *Nepal small area estimation of poverty 2011 (English)*. Nepal small area estimation of poverty 2011; volume 1. Washington DC: World Bank.

Unmet need for family planning



UNFPA Country Capacity Strengthening Workshops

UNFPA carries out regional SAE capacity strengthening workshops with the objective of enhancing the capability of relevant government departments/ ministries and other counterparts to:

- Assess and harmonize census and survey data for use in small area estimation;
- Develop predictive models and select the "best" fitting models for the indicators of interest;
- Generate reliable small-area-level estimates for the key indicators using the ELL method by combining census and survey data.

Many Possible Applications of SAE

Apply the SAE methodology to other development issues such as antenatal care, skilled birth attendance, harmful practices, etc.

Use SAE results for targeting locations of the most vulnerable groups. e.g. clusters of high poverty areas.

Linking SAE results with other geospatial data for improved programming, e.g. access to schools and health services, exposure in climate-related hazard zone, etc.

Proportion of demand for modern contraception satisfied
SDG INDICATOR 3.7.1