

# DYNAMIS-POP NEPAL

## Dynamic Micro-Simulation for Population and Development Projections

Discussion Note

Martin Spielauer ([martin.spielauer@dms-c.com](mailto:martin.spielauer@dms-c.com)) and Olivier Dupriez ([odupriez@worldbank.org](mailto:odupriez@worldbank.org))

November 2018

This note gives a short overview of the DYNAMIS-POP model and of its potential applications in Nepal. DYNAMIS-POP is a portable dynamic micro-simulation model. It is designed to create and run realistic data-driven simulations with emphasis on applications addressing development issues. It is based on data readily available for most developing countries. As a highly modular system it can be adapted and refined for specific uses in a broad field of policy-relevant applications. The model is built on free and open source software applications.

## Dynamic Microsimulation

Micro-simulation in the context of socioeconomic applications can be perceived as an experiment with a virtual society of thousands—or millions—of individuals. Central to dynamic micro-simulation is the explicit modeling of the time dimension, following people and their families or households over time, rather than performing a before-and-after comparison. Dynamic modeling lends itself naturally to the modeling of policies with a longitudinal component, e.g., educational investments, especially in the context of general rapid social, economic, and demographic change that make it difficult to assess the contribution of individual policies to overall trends without tracking and comparing the lives of individuals who form a society.

DYNAMIS-POP is implemented as a user-friendly Windows application with a graphical user interface for parameter handling, scenario generation and table output. Simulation output can also be visualized (e.g. maps, population pyramids) using a set of documented R scripts.

The model (including statistical analysis and technical implementation) is fully documented and reproducible. All software components are freely available. The official website will be launched soon, a test version is available at [www.dms-c.com/dynamis](http://www.dms-c.com/dynamis).

## Adaptation to Nepal

Two preliminary versions of DYNAMIS-POP for Nepal are available, one based on the 2001 Census and DHS data, the other one based on the 2011 Census and DHS data. Starting simulations in 2001 allows for retrospective validations.

## Scope and Modules

DYNAMIS-POP is designed as a modeling platform which can be developed in various directions. The current focus lies on detailed regional population projections, child mortality, and education projections as well as on the effects of educational improvements on fertility and child mortality. Recent and ongoing additions address stunting, human capital, and vaccination.

The model complements existing macro projection models by depicting the population and its development in its heterogeneity and the creation of realistic individual life courses. At the same time, DYNAMIS-POP can replicate existing population projections on the aggregate level.

DYNAMIS-POP is highly modular. The following modules comprise the core of DYNAMIS-POP:

- **Fertility:** we use models estimated from Census (first birth) and DHS (higher order births) data which, besides age, account for parity, education, marital status and birth intervals. The number of births can be aligned to macro projections, thus does not differ from available population projections. While producing the same number of births as in available projections, babies are realistically distributed to women which allows using mothers' characteristics when modeling other processes.
- **Mortality:** we use a standard UN life table and existing projections of male and female life expectancy. For child mortality (age 0-4) an optional module can be used based on DHS data. In this case, the results can be aligned to the macro projection for the first year, but then can differ from macro projections as we account for composition effects by mothers' characteristics (teenage pregnancies, education) which influence child mortality.
- **Migration:** For modeling immigration, emigration and internal migration we use period rates calculated from the Census data accounting for age and sex. DYNAMIS-POP allows three geographical levels: national – regional – provincial.
- **Primary Education:** We use a cohort model based on the Census to decide about school entry and graduation by province. The model allows to continue recent trends or easily create alternative scenarios. The model also allows to model the intergenerational transmission of education by mothers' or other individual characteristics (ethnicity, stunting, etc.).
- **School grade tracking:** based on parameters on grade repetition, dropout and the distribution of the highest attained grade of dropouts we generate realistic school histories which can be used for modeling required numbers of teachers and classrooms etc.
- **Secondary education:** is based on projected period intake, repetition, and dropout rates.
- **First Marriage** is modeled fitting a Coale McNeil model by education based on Census data and – in the base scenario - trends informed by this analysis.
- **Ethnicity:** We model the intergenerational transmission of castes but at this step do not use this information in any process.

## Required Data and Data Issues

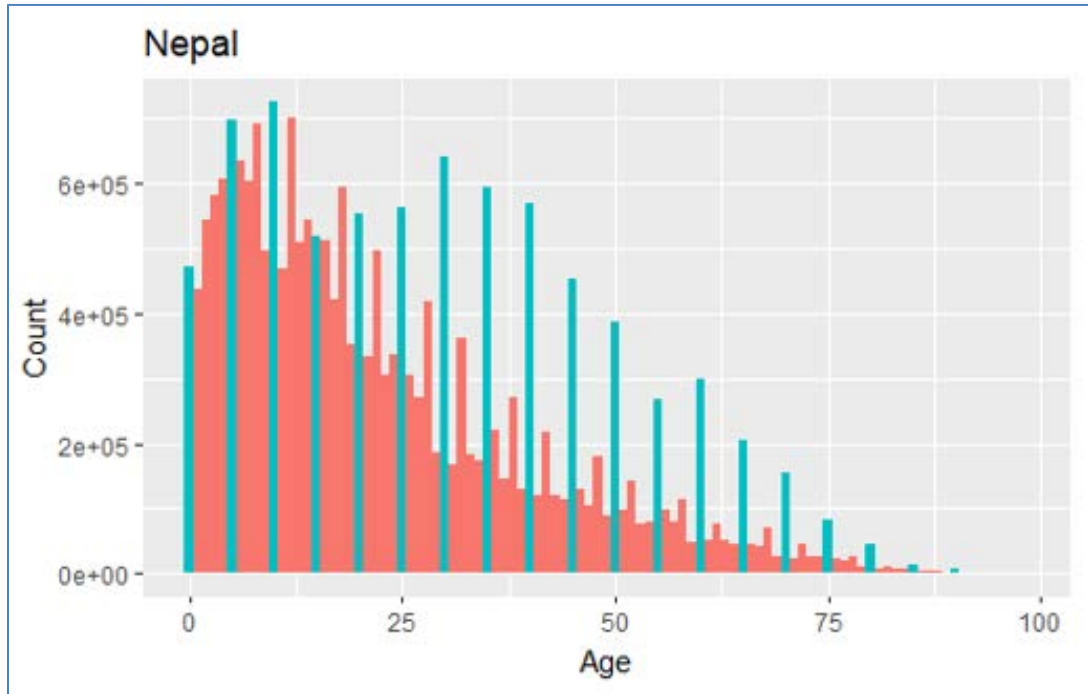
### Required data

The starting population of DYNAMIS-POP as well as the model parameters are based on Census and DHS data, which are available for Nepal for 2001 and 2011. All required data preparation and analysis work is highly automated, and all scripts (R code) are documented and available online.

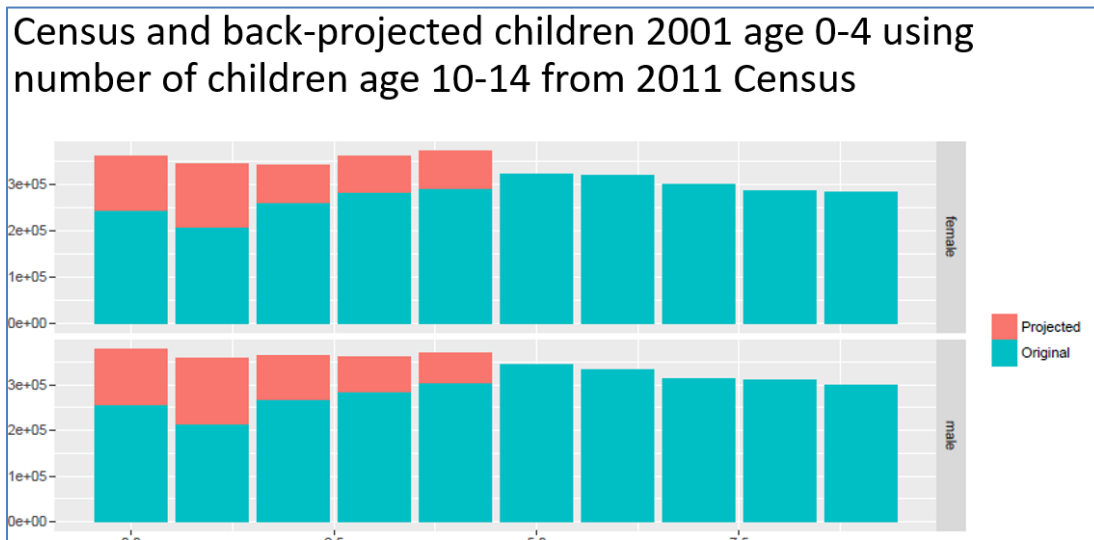
### Data issues

The quality of DYNAMIS-POP projections depends on the quality of data; a suite of statistical tools (R simPop package) addressing common data issues was developed in a complementary project. In the case of Nepal, detected – and treated -- issues include age heaping and the under-reporting of children below age 5.

The under-reporting of young children in the Census is noticeable in a comparison of DHS with Census data. It is confirmed by back-projecting to 2001 the number of 10-14-year-old children born in Nepal reported in the 2011 Census (accounting for mortality).



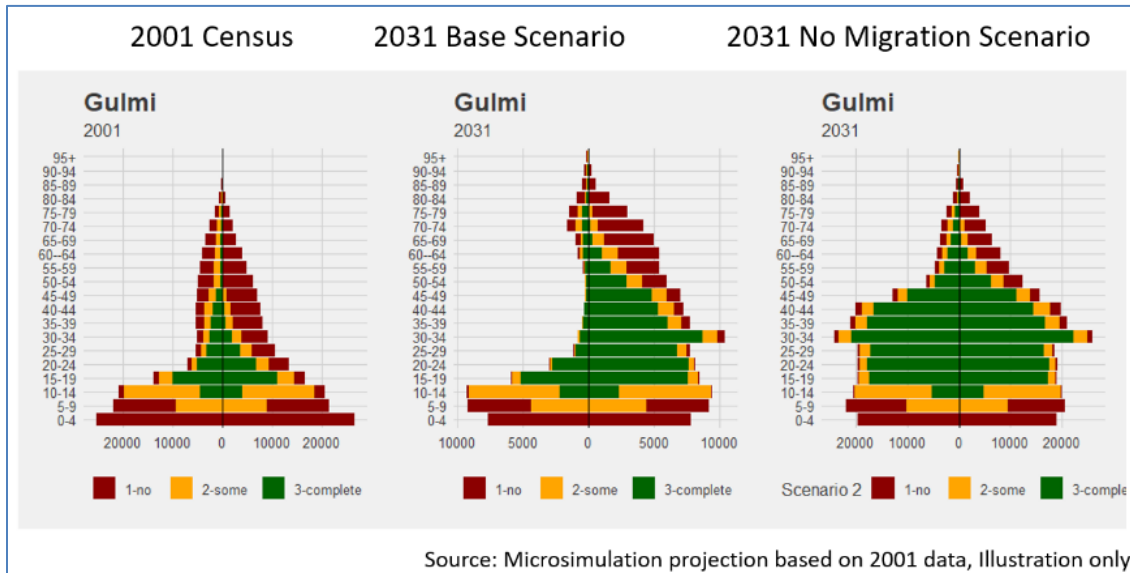
*Illustration: Age heaping in the 2001 Census*



*Illustration: Under-reporting of children 0-4 in the 2001 Census*

## Scenarios based on period observations

Another issue affecting population projections in Nepal is the (non-) treatment of expected return migration of labor migrants which cannot be captured from period data alone and – in a mechanical projection based on current emigration and immigration patterns – would lead to a depopulation (males) in some provinces.



*Illustration: Population projection for Gulmi based on 2001 migration rates*

As a micro-simulation model, DYNAMIS-POP can track individual migration careers. The model could be easily extended for allowing scenario creation on expected return migration.

## Application Fields

DYNAMIS-POP is suited for

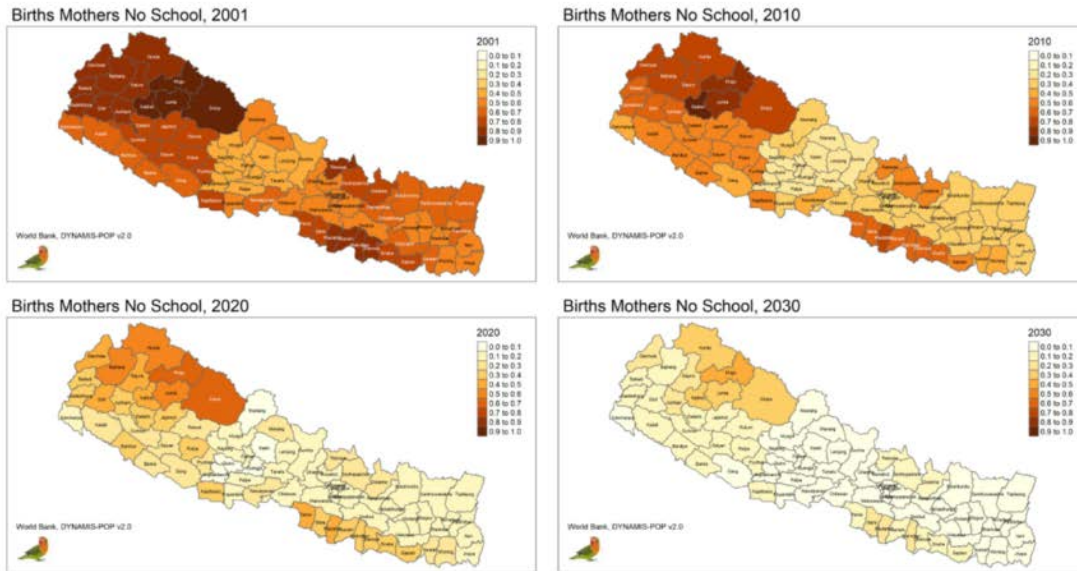
- Regional population projections, with disaggregation by various individual characteristics like ethnicity (caste) or education.
- Modeling of realistic scenarios of return migration.
- Detailed education projections include the tracking of students by grade (school planning) and the modeling of educational differences between population groups. Children out of school and school dropout.
- Child mortality
- Gender issues (teenage pregnancies, etc.)

This list of possible applications is being expanded. We currently focusing on health (vaccination, stunting) and are exploring modeling of human capital (as a function of stunting, mortality risks and education).

The model allows studying downstream-effects (e.g. the effect of educational improvements on fertility, child mortality) and decomposition of changes to underlying factors (e.g. quantifying composition effects).

## Examples

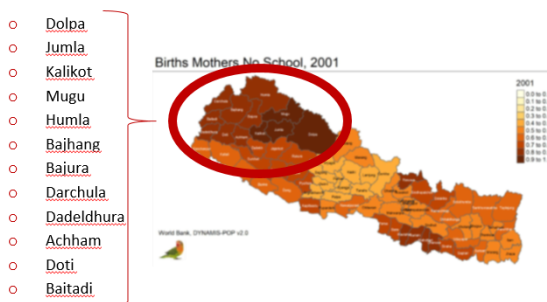
### Example: Births (%) by mothers never in school



Source: Microsimulation projection based on 2001 data, Illustration only

### Example: Infant (0-4) deaths 2015-35

- **Base Scenario:** Education following current trend
- **Alternative Scenario:** Universal primary education for all born 2001+



	Child Deaths	
	2016-25	2026-35
Base Scenario - North-West	11930	7730
Universal Primary Scenario - North-West	11200	5550
	-6.1%	-28.2%
<hr/>		
	2016-25	2026-35
Base Scenario - Nepal	148960	101290
Universal Primary Scenario - Nepal	143560	86840
	-3.6%	-14.3%

Source: Micro-simulation projection based on 2001 data, Illustration only.  
 Validation: UNICEF 24,000 child deaths in 2012, the projected number in the micro-simulation is 21,230 for 2012

## DYNAMIS-POP in the Context of SDGs

DYNAMIS-POP is highly modular and can serve as a base for model extensions in various dimensions. In the medium term, this opens a wide range of opportunities to develop the model to a tool supporting the projection and analysis of processes of sustainable development and related indicators.

### Currently available indicators

Different to macro population projections, DYNAMIS-POP has specific modules addressing child mortality by mother's characteristics, first unions by education, and births by union status, education, and parity. In this respect, three SDG indicators can not only be projected, but the model can be used for analysis of policy effects (e.g. educational expansion, marriage age restrictions) on demographic outcomes.

**Goal 3.** Ensure healthy lives and promote well-being for all at all ages

- C030201 Under-five mortality rate
- C030702 Adolescent birth rate (aged 10-14 years; aged 15-19 years) per 1,000 women in that age group

**Goal 5.** Achieve gender equality and empower all women and girls

- C050301 Proportion of women aged 20-24 years who were married or in a union before age 15 and before age 18

### Indicators which could be added easily to the current model framework

Potential extensions within the existing model framework mostly concern mortality, education, and health.

**Goal 2.** End hunger, achieve food security and improved nutrition and promote sustainable agriculture

- C020201 Prevalence of stunting (height for age <-2 standard deviation from the median of the World Health Organization (WHO) Child Growth Standards) among children under 5 years of age. A base version of such a module is available for Nepal already

**Goal 3.** Ensure healthy lives and promote well-being for all at all ages

Child mortality is currently modeled by years of age; singling out neonatal deaths (defined as a death during the first 28 days of life) by applying separate mortality rates for this early phase of life is a possible refinement. Also, a module for maternal mortality could be added.

- C030101 Maternal mortality ratio
- C030202 Neonatal mortality rate

**Goal 4.** Ensure inclusive and equitable quality education and promote lifelong learning opportunities

A second set of rather direct model extensions concern education indicators. A specific module of early childhood education (with downstream effects on future educational careers) could be added.

- C040202 Participation rate in organized learning (one year before the official primary entry age)

Microsimulation projections allow to generate general participation rates (at least in formal education) and support the disaggregation of all indicators to specific population groups, at least covering parts of the required dis-aggregations.

- C040301 Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex

- C040501 Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated

While the model so far allows tracking students through the formal school system by grade, we could also add a module which links school attendance and careers with proficiency outcomes.

- C040101 Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex
- C040601 Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex

Extension of individual characteristics allowing further dis-aggregations and addressing specific groups

- Ethnicity: (Casts) this is modeled already for Nepal, but the variable is not used yet.
- Disability: People with disabilities are addressed specifically in various indicators. A module on disability status could be added

## Potential 'non-economic' extensions and applications

**Goal 3:** Ensure healthy lives and promote well-being for all at all ages

### (a) Adding Health Infrastructure

Adding Health Infrastructure information is a possible road for future model development, with possible interactions with internal migration (by attractiveness of province/region) and child and maternal mortality.

- C030c01 Health worker density and distribution
- C030801 Coverage of essential health services

Some outcomes which to a large extent depend on infrastructure could be modeled specifically

- C030102 Proportion of births attended by skilled health personnel
- C030b01 Proportion of the target population covered by all vaccines included in their national program. A module on child vaccination is already under development.

### (b) Modeling of diseases and cause-specific mortality

Modeling of cause-specific mortality as well as the spread of diseases are common microsimulation applications. Depending on national priorities, DYNAMIS might be used as base of the modeling of specific diseases (e.g. Malaria, HIV) and interventions.

- C030301 Number of new HIV infections per 1,000 uninfected population, by sex, age and key populations
- C030302 Tuberculosis incidence per 100,000 population
- C030303 Malaria incidence per 1,000 population
- C030304 Hepatitis B incidence per 100,000 population
- C030305 Number of people requiring interventions against neglected tropical diseases
- C030401 Mortality rate attributed to cardiovascular disease, cancer, diabetes or chronic respiratory disease

Incidence rates of diseases typically have an age profile and regional variation, this might make even simple models concentrating on population projections by few characteristics useful for projecting future rates

### (c) Health-related behaviors

Another typical application field of microsimulation for which DYNAMIS could serve as base are health-related behaviors like alcohol and tobacco consumption.

- C030502 Harmful use of alcohol, defined according to the national context as alcohol per capita consumption (aged 15 years and older) within a calendar year in litres of pure alcohol
- C030a01 Age-standardized prevalence of current tobacco use among persons aged 15 years +

Specialized model applications are also thinkable for unsafe water/sanitation and related diseases and mortality. Such an application would draw on the ability of microsimulation to combine infrastructure data, fine-grained geography, and individual behavior.

- C060101 Proportion of population using safely managed drinking water services
- C060201 Proportion of population using safely managed sanitation services, including a hand-washing facility with soap and water
- C030902 Mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene (exposure to unsafe Water, Sanitation and Hygiene for All (WASH) services)

## Economic modeling

This involves the modeling of labor force participation, employment, income and wealth. Such extensions require extensive planning and research but would extend the range of applications to a wide range of SDG indicators. Entry point might be cross-sectional imputation models.

### **Goal 1.** End poverty in all its forms everywhere

- C010101 Proportion of population below the international poverty line, by sex, age, employment status and geographical location (urban/rural)
- C010201 Proportion of population living below the national poverty line, by sex and age
- C010202 Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions
- C010301 Proportion of population covered by social protection floors/systems, by sex, distinguishing children, unemployed persons, older persons, persons with disabilities, pregnant women, newborns, work-injury victims and the poor and the vulnerable
- C010401 Proportion of population living in households with access to basic services

### **Goal 2.** End hunger, achieve food security and improved nutrition and promote sustainable agriculture

- C020101 Prevalence of undernourishment
- C020102 Prevalence of moderate or severe food insecurity in the population, based on the Food Insecurity Experience Scale (FIES)
- C020202 Prevalence of malnutrition (weight for height  $>+2$  or  $<-2$  standard deviation from the median of the WHO Child Growth Standards) among children under 5 years of age, by type (wasting and overweight)
- C030802 Proportion of population with large household expenditures on health as a share of total household expenditure or income

### **Goal 8.** Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all

- C080601 Proportion of youth (aged 15-24 years) not in education, employment or training
- C080701 Proportion and number of children aged 5-17 years engaged in child labor, by sex and age



**Goal 10.** Reduce inequality within and among countries

- C100201 Proportion of people living below 50 per cent of median income, by sex, age and persons with disabilities