DYNAMIS-POP

November 2018

martin.spielauer@dms-c.com

DYNAMIS-POP Background

A portable dynamic socio-demographic micro-simulation platform for developing countries

- Based on micro-data readily available in most countries: Census + DHS or MICS
- Portable: so far Mauritania 2013, Nepal 2001, Nepal 2011
- Main focus (to date, -POP) on detailed population projections, complementing available national and regional projections by adding information on education careers, family demographics, ethnicity, health
- Ability to reproduce existing aggregate projections, but adding geographic and life-course detail, modeling in family and regional context
- Modular platform, extendable for applications in a variety of policy-relevant fields

DYNAMIS-POP Philosophy

- Maximum automation of workflow
  - Automated generation of model parameters (from standardized files)
  - Most simulation code generic
  - Scripts for ex-post analysis and visualization
- Reproducible
  - Detailed documentation incl. step-by-step analysis and implementation guide
  - All software components freely available for download
- User friendly: graphical user interface (GUI) and intuitive parameters
- Rich Output
  - Output tables (exportable: incl. coefficients of variation of each table cell)
  - Micro-data output (cross-sectional panel data; individual histories)

Current modules

- Demographic core reproducing a cohort-component model
  - Fertility
  - Mortality
  - Migration: immigration, emigration, internal migration
- Other core modules going beyond macro projections
  - Primary education ‚fate‘
  - Transmission of ethnicity
  - First marriage
- Refined and optional modules
  - Educational transmission
  - Refined Fertility by parity, education, marital status, time since last birth
  - Child mortality by mother‘s characteristics
  - Primary education tracking: following students through grade system
  - School planning: required classrooms, teachers etc.
  - Secondary education
  - Stunting + HCI

Modules: Fertility

- Base Version
  - Age-specific fertility distribution by year
  - Total Fertility Rate (TFR) by year
- Extended Version
  - First births by age, union status, education, province
  - Higher order births by education, time since last birth
  - Separate trends by birth order
- Alignment: forcing the model to reproduce aggregate outcomes while respecting relative fertility differences thereby generating realistic life-courses. Choices:
  - Not aligned
  - Aligned to total births of base version (same number of births)
  - Aligned to total births by age of base version (same age-specific fertility rates)

Example: Births (%) by mothers never in school

Births Mothers No School, 2001

Births Mothers No School, 2010

Births Mothers No School, 2020

Births Mothers No School, 2030

Source: Microsimulation projection based on 2001 data, Illustration only

Modules: Mortality

- Base Version
  - Standard life table of age-specific rates by sex
  - Life expectancy by calendar year and sex
- Refined child mortality model (ages 0-4)
  - Age baseline
  - Relative risks by mothers education and age group
  - Age-specific overall trends
- Alignment options (refined model)
  - Without
  - Initial alignment to base model – trends from base
  - Initial alignment to base model – specific trends

Example: Child (0-4) deaths 2015-35

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

- Dolpa
- Jumla
- Kalikot
- Mugu
- Humla
- Bajhang
- Bajura
- Darchula
- Dadeldhura
- Achham
- Doti
- Baitadi

	Child Deaths	
	2016-25	2026-35
Base Scenario - North-West	11930	7730
Universal Primary Scenario - North-West	11200	5550
	-6.1%	-28.2%

	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

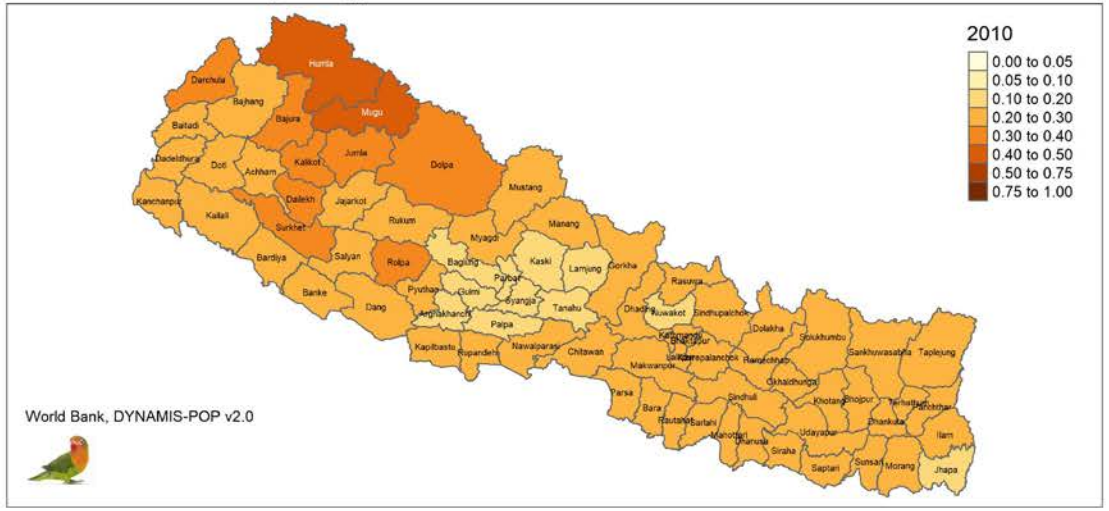
Modules: Education

- Base Version
  - Probability to enter and graduate from primary education by sex, year of birth, district. (typically modeled by logistic regression containing a logarithmic trend)
  - Period model for secondary education (parameterized by intake, progression, repetition, dropout rates as available e.g. by UNESCO)
- Refinements
  - Education transmission by mother's education + effect of stunting (odds ratios; outcomes can be aligned for one or all years)
  - Students tracked through school system by grade (using intake, progression, repetition, dropout information (e.g. from UNESCO) aligned to modeled outcomes)
  - School resource planning of required classrooms and teachers: Target path for classroom sizes and teacher/student ratios

Example: Children 9-11 out of school

School Absence (9-11), 2001

School Absence (9-11), 2010



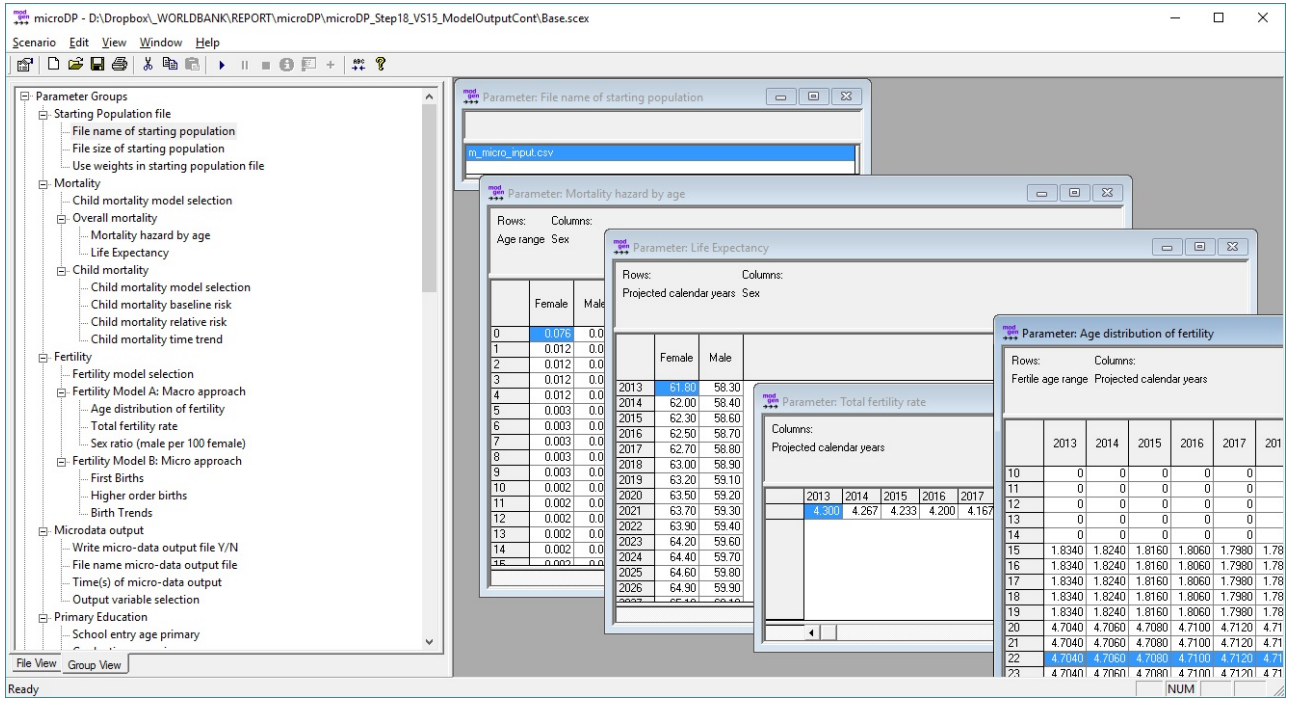
School Absence (9-11), 2020

School Absence (9-11), 2030

Source: Microsimulation projection based on 2001 data, Illustration only

Implementation

- Implemented in Modgen (Statistics Canada), a generic microsimulation programming language based on C++
- Graphical User Interface
- Scenario support
- Rich, exportable table output
- Various table views: values, coefficient of variation
- Fully documented (Help files for user interface and model)
- Fast (can simulate millions of interacting agents on a standard PC)



Work-Flow – Creation of a new country version

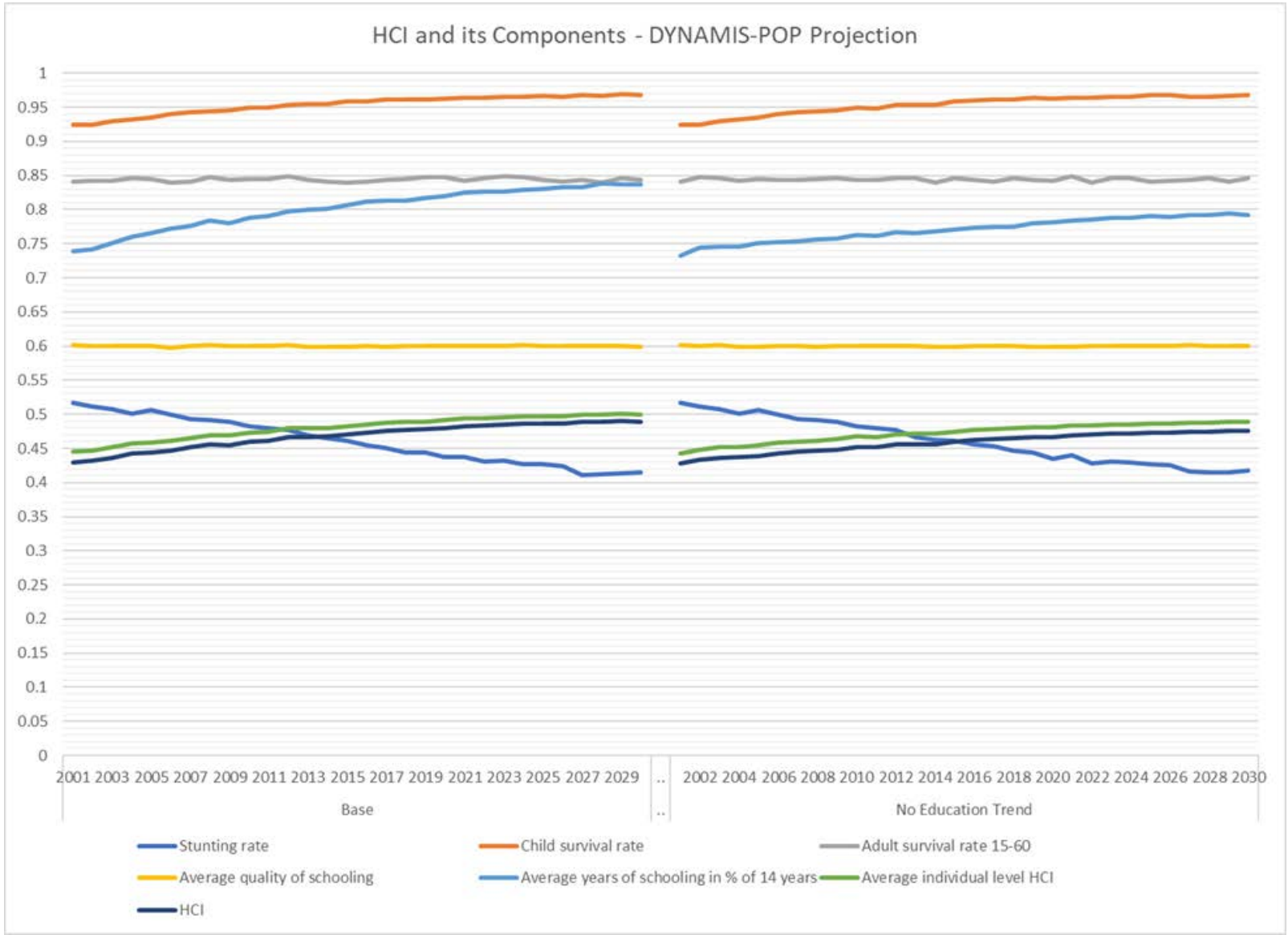
- Data preparation: creation of 4 standardized micro-data files. Some other files: macro projections, shape files for map output
- Country-specific R setup script: file names and locations and calendar time values as models might start at different start years.
- Run R input analysis scripts: (currently 16 numbered scripts) for parameter estimation, production of all parameter files and a the starting population.
- Country specific simulation code file: one (of the currently 33) code files (modules) is country specific: name of districts, mapping to regions, start year, etc.
- Compile and start the new model

HCI Index (Demo, Nepal, projected from 2001)

- Module for stunting: stunting rates by sex and mother's education from DHS (projects composition effects only, no trends)
- Preschool module: ad-hoc
- Module for HCI: Output of all components, aggregated HCI and average individual index

- General mortality: period rates frozen from 2018 onwards
- Child mortality by mother's age and education
- Primary school: cohort model by sex, mother's education, stunting, region, trend
- Secondary: time-invariant take-up, repetition, progression rates
- School quality: current national average

HCI Index (Demo, Nepal, projected from 2001)



What DYNAMIS can add: (1) cohort studies

- HCI Projections: retrospective, prospective
- Benchmark projections: helping to assess policy effects
  - Status quo on individual level: how would HCI change if nothing changes for given individual parental, ethnical, regional… background.
  - How would HCI change if existing population projections are accurate?
- Downstream effects / what-if scenarios: e.g. effect universal primary schooling
- Regional disaggregation
- Decomposition of changes
  - Impact of changes in component (e.g. child mortality improvements)
  - Decomposition of changes within components (e.g. composition versus other effects)

What DYNAMIS can add: (2) population studies

- Projections of the human capital of the (e.g. working age) population
  - Imputation of human capital to current population of all ages
  - Different perspectives: human capital of population alive
- Economic modeling
  - Production functions require input of human capital of active population
  - Modeling of labor force participation by individual characteristics
- What-if / policy scenarios from population perspective
  - E.g. How would educational improvements in specific population groups impact the future labor force participation and human capital
  - What is the timeline of such changes

Supplementary information

Data requirements

- Data requirements met for most countries by:
  - A population Census
  - Survey data on demographic events:
    - MICS: Multiple Indicators Cluster Surveys (UNICEF)
    - DHS: Demographic Health Survey
- Four essential data files:
  - Residents
  - Recent emigrants
  - Children
  - Birth histories

```
M_WEIGHT    Weight (123.456)
M_AGE       Age (in years, 16.789)
M_MALE      Sex (female 0, male 1)
M_DOB       District of birth (0..m, m = abroad)
M_DOR       District of residence (0..n)
M_PDIST     District 12 months ago (0..m, m = abroad)
M_EDUC      Primary education (0 non, 1 some, 2 graduated)
M_PARITY    Parity (0, 1..)
M_BIR12     Number of births past 12 months (0, 1, 2)
M_AGEMAR    Age at first marriage (in years, 16.789, 999 never married)
M_AGEBIR    Age at most recent birth (in years, 16.789, 999 childless)
M_ROB       Region of birth (0..b, b = abroad)
M_ROR       Region of residence (0..a)
M_PREG      Region 12 months ago (0..b, b = abroad)
M_ETHNO     Ethnicity (0..y)
```

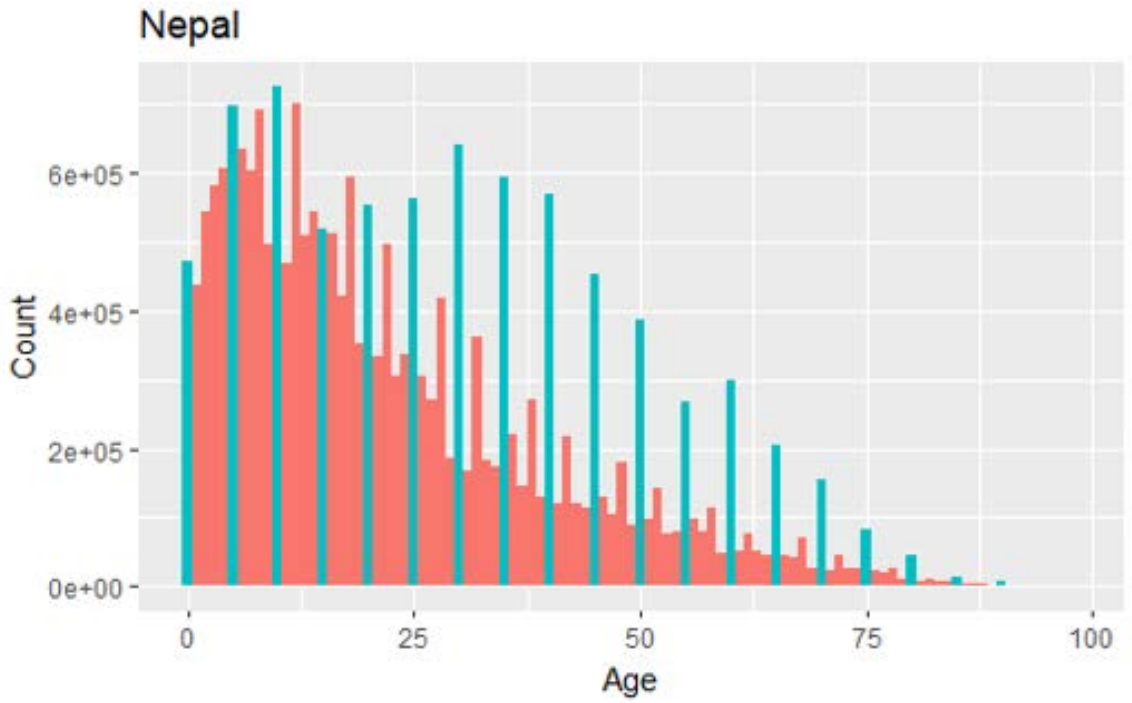
```
M_WEIGHT    Weight (123.456)
M_PDIST     District 12 months ago (0..n)
M_PREG      Region 12 months ago (0..x)
M_AGE       Age (in years 18.901)
M_MALE      Sex (0 female, 1 male)
```

```
M_BIRTH     Birth of child (month since 1900)
M_DEATH     Death of child (month since 1900; 9999 if alive)
M_MALE      Sex of child ( 0 female, 1 male)
M_WEIGHT    Weight (123.456)
M_AGEMO     Age group of mother when giving birth (months)
M_EDUCMO    Primary education of mother (0 non, 1 some, 2 graduated)
M_INTERV    Time of interview (month since 1900)
```

```
M_B01       Month of 1st birth (months since 1900; 9999 for non)
...
M_B14       Month of 14th birth (months since 1900)
M_WEIGHT    Weight (123.456)
M_BIRTH     Birth (months since 1900)
M_EDUC      Primary education (0 non, 1 some, 2 graduated)
M_REG       Region of residence (0..n)
M_INTERV    Time of interview (months since 1900)
M_MAR       Time of first marriage (months since 1900; 9999 never married)
```

Data Issues

- Complementary project and R packages for addressing typical data issues and for synthetic population generation
  - Age Heaping
  - Under-reporting of children
  - Imputation of missing variables
  - Generation of synthetic datasets



Census and back-projected children 2001 age 0-4 using number of children age 10-14 from 2011 Census