## **Sub-national mapping of population pyramids and dependency ratios in Africa and Asia**

Keywords*: population age and sex structure; age-dependency; Asia; Africa; sub-national maps*

**Authors**

Carla Pezzulo1,2, Graeme M. Hornby1,3, Alessandro Sorichetta1,2, Andrea E. Gaughan4, Catherine Linard5,6, Tomas J. Bird1,2, David Kerr1, Christopher T. Lloyd1, and Andrew J. Tatem1,2

**Affiliations**

1. WorldPop, Department of Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK

2. Flowminder Foundation, Roslagsgatan 17 SE-11355, Stockholm. Sweden

3. GeoData, University of Southampton, Southampton SO17 1BJ, United Kingdom;

4. Department of Geography and Geosciences, University of Louisville, Louisville, KY 40292;

5. Department of Geography, University of Namur, B-5000 Namur, Belgium;

6. Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, B-1050 Brussels, Belgium;

Corresponding author: Carla Pezzulo ([C.Pezzulo@soton.ac.uk](mailto:C.Pezzulo@soton.ac.uk))

**Abstract**

The age group composition of populations varies substantially across continents and within countries, and is linked to levels of development, health status and poverty. The subnational variability in the shape of the population pyramid as well as the respective dependency ratio are reflective of the different levels of development of a country and are drivers for a country’s economic prospects and health burdens. Whether measured as the ratio between those of working age and those young and old who are dependent upon them, or through separate young and old-age metrics, dependency ratios are often highly heterogeneous between and within countries.

Assessments of subnational dependency ratio and age structure patterns have been undertaken for specific countries and across high income regions, but to a lesser extent across the low income regions. In the framework of the WorldPop Project, through the assembly of over 100 million records across 6,389 subnational administrative units, subnational dependency ratio and high resolution gridded age/sex group datasets were produced for 87 countries in Africa and Asia.

**Background & Summary**

Populations in low income countries are generally younger and growing more rapidly compared to those in high income countries. Africa has the highest rate of population growth and more than half of the global population growth between now and 2050 is expected to occur in Africa1. Asia is projected to be the second largest contributor to future global population growth, adding 0.9 billion people between 2015 and 20501.

Population demographics have been central to the Millennium Development Goals and are key to the Post-2015 Sustainable Development Agenda2, and development policies need to account for population dynamics and their relationships with social, economic and environmental factors3,4. One aspect characterising low income countries is the higher proportion of children relative to working age individuals5,6. Moreover, given the increasing number of elders dependent on working age individuals, population aging is also an emerging issue in some low and middle income countries7,8,. The ratio between the share of the ‘dependent age’ population (either <15 years or >65 years) to the remaining ‘working age’ population is generally quantified through the ‘age dependency ratio’5,9,10. A declining share of children and/or elders and an increase in the working-age adult population lowers the dependency ratio and can lead to opportunities for development11,12. For example, households with a lower number of dependents tend to increase the share of savings and to invest in the human capital of their children13,14, and countries experiencing a rise in the number of working-age people can expand their productivity15. In contrast, experiencing high young-age and old-age dependency ratios means that working age populations face a greater burden in supporting and providing the social services needed by children and elders16. In areas such as the sub-Saharan Africa and South Asia, age dependency ratios remain relatively high, and only in some cases have they started to slowly decline, prompting opportunities for development and economic growth17,18, with the potential to enter into the so-called ‘demographic dividend’19.

Age-sex pyramids and dependency ratios are generally measured, mapped and summarised at national levels, in spite of the fact that large subnational demographic heterogeneities exist within countries. Studies have identified and explored the role of subnational or household-level dependency ratios in some countries20-22, but these have focussed mostly on high income regions23. However, subnational demographic data are becoming increasingly available for countries in low income regions24 and this trend is being encouraged for highlighting inequalities, targeting populations at risk24, and guiding the development of targeted policies for the Sustainable Development Goals25.

Contemporary and spatially detailed datasets that identify areas where policy relevant age groups (e.g., children under the age of five, young persons, women of childbearing age and the elderly) reside are a fundamental prerequisite for monitoring demographic changes15, supporting geographic targeting of policies ([www.undatarevolution.org](http://www.undatarevolution.org)), planning vaccination interventions26), identifying age groups at risk of mortality and morbidity27 and allocating resources effectively28,29. Accurately depicting the distribution of young age and old age dependency ratios at subnational levels aids in identifying opportunities for economic growth18 and possible gains from the demographic dividend30. It also helps in investigating other socio-economic conditions usually related to high dependency ratios, such as high fertility rates31, low levels of child enrolment in schools32 and low opportunities for productive employment 33.

In the framework of the WorldPop Project (www.worldpop.org)34, advancing the approach described in Tatem et al.24, an open access archive of (i) gridded subnational dependency ratio datasets and (ii) high resolution gridded 5-year age/sex group count datasets, for 87 countries in Africa and Asia (excluding Middle Eastern countries and Japan), has been created using subnational data on population age and sex structures for circa 2010 and the corresponding WorldPop gridded continental population count datasets (Fig. 1). In particular, over 100 million records were extracted from census data, census microdata and household survey data, assembled and used to derive the proportional age and sex structures for 6,389 subnational administrative units.

**Methods**

Subnational tabular data on population age and sex structures were gathered on a country by country basis. The table format and structure varied significantly from one country to the next, reflecting the wide range of data sources utilised in this project. For example, some countries presented absolute counts per age group whilst others presented proportions. The thresholds that determined the age groups also varied from country to country, ranging from 1 to 5 years. In addition, some datasets represented the full population (e.g. national census data), whilst others were a sub-sample of the population (e.g. micro-census, household surveys). These challenges required a process of standardisation aimed at harmonizing these disparate data sources into a set of consistent tables in which each row represents a subnational administrative unit and the fields contain the corresponding proportion of males, females and people in each 5-year age group (from aged 0 to 4 up to a final group aged 65 and over).Standardized country data tables were subsequently joined to their respective subnational spatial boundaries and the spatial datasets generated were then merged together to produce two continental vector datasets (i.e. Africa and Asia) depicting the proportional age and sex structure in each subnational administrative unit.

Subnational proportional age structures were then used to calculate combined, young-age and old-age dependency ratios for each administrative unit in Africa and Asia. According to the standard definitions most widely found in the literature10,35, the combined dependency ratio was measured as the ratio of dependents younger than 15 and aged 65 and over to the population aged 15-64; the young-age dependency ratio as the ratio of dependents younger than 15 to the working-age population aged 15-64; and the old-age dependency ratio as the ratio of dependents aged 65 and over to the working-age population. Eventually, all vector subnational dependency ratio datasets were rasterized at a resolution of 30 arc seconds (approximately 1 km at the equator).

A series of gridded 5-year age/sex group count datasets for Africa and Asia, with a spatial resolution of 30 arc seconds, were generated by gridding the subnational proportional age and sex structure datasets and overlying them with the corresponding WorldPop gridded continental population count dataset (http://www.worldpop.org.uk/data/data\_sources/); with the latter adjusted to match United Nations Population Division (UNPD) estimates for 201036.

**Data Collection**

Data on population age and sex structure, as close as possible to 2010, were collected at the finest spatial level available for all African and Asian countries listed in Table 1 and 2. Where data for a given country were available from multiple sources, in addition to the year in which the data were collected, another key determinant, such as the sample size of the data, was considered in order to decide which data should be used.

The most reliable, complete and accurate source for subnational population composition is usually represented by country-level census-based data that cover the whole country population. Thus, for those countries where age and sex structure data were available from a recent national census, these were collected at the highest administrative level available, along with their corresponding spatial boundaries.

For those countries where recent full census data were not available, alternative sources were sought. Priority was given to census microdata obtained from the Integrated Public Use Microdata Series International (IPUMSI) database37. Census microdata represent household-level records derived from census data by sampling a representative fraction of the population (with sample size generally between 2 and 15% of the full census).

Where neither full census nor census microdata close to 2010 were available, national household survey data were obtained from the Demographic and Health Surveys (DHS), DHS special Malaria Indicators Surveys (MIS) and Aids Indicator Surveys (AIS)38, Social Indicator Surveys (SIS)39 or Multiple Indicator Cluster Surveys (MICS)40. Whilst providing a representation of the population closer to 2010, compared to outdated full census and census microdata, household surveys tend to suffer from a constrained sampling framework, with sample sizes being typically less than 1% of the national population and full range extending from 0.04% to 12%. Nevertheless, household survey data are designed to be representative at both national and subnational levels (typically administrative level 1, generally equivalent to provinces) and thus can be used to derive the corresponding age and sex structures. Household surveys were prioritised by survey year and sample size.

Tables 1 and 2 list data sources on a country-by-country basis for Africa and Asia, respectively, and provide information on the year of data collection and the subnational administrative level at which they were collected. In summary, the data for 30 of the 87 countries were derived from full national population and housing censuses, 17 from IPUMSI census microdata, 19 from traditional DHS, 4 from DHS-MIS/AIS/SIS and 8 from MICS (Tables 1 and 2).

For the remaining 9 countries where subnational data on population age and sex structures were not available (namely Libya, Eritrea, Western Sahara, Equatorial Guinea, Brunei, Myanmar, Papua New Guinea, Sri Lanka and Turkmenistan) 2010 UNPD country level estimates35 were used.

| **Country** | **Data type used** | **Year** | **Administrative unit level** | **Data source** |
| --- | --- | --- | --- | --- |
| Algeria | Census | 2008 | 1 | Office National des Statistique, Algeria |
| Angola | MIS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Benin | DHS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Botswana | Census | 2006 | 2 | Central Statistics Office, Botswana |
| Burkina Faso | Census | 2006 | 1 | Institut National de la Statistique et de la Demographie (INSD), Burkina Faso |
| Burundi | DHS | 2010 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Cameroon | DHS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Central African Republic | MICS | 2006 | 1 | UNICEF |
| Chad | DHS | 2004 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Congo | AIS | 2009 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Congo, Democratic Republic | DHS | 2013 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Cote d'Ivoire | AIS | 2005 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Djibouti | MICS | 2006 | 1 | UNICEF |
| Egypt | Census microdata | 2006 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Equatorial Guinea | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Eritrea | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Ethiopia | Census | 2007 | 1 | Central Statistical Agency of Ethiopia |
| Gabon | DHS | 2012 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Gambia | MICS | 2006 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Ghana | Census | 2010 | 2 | United States Census Bureau, USAID |
| Guinea | DHS | 2012 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Guinea-Bissau | MICS | 2006 | 1 | UNICEF |
| Kenya | Census | 2010 | 1 | United States Census Bureau, USAID |
| Lesotho | Census | 2004 | 1 | Lesotho Bureau of Statistics |
| Liberia | Census | 2008 | 1 | Liberia Institute of Statistics and Geo-Information Service |
| Libya | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Madagascar | DHS | 2009 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Malawi | Census | 2008 | 2 | National Statistical Office, Malawi |
| Mali | DHS | 2012 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Mauritania | MICS | 2007 | 1 | UNICEF |
| Morocco | Census | 2004 | 2 | Haut Commissariat au Plan, Morocco |
| Mozambique | Census | 2007 | 2 | Instituto Nacional de Estatistica, Mozambique |
| Namibia | Census | 2010 | 2 | United States Census Bureau, USAID |
| Niger | DHS | 2012 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Nigeria | Census | 2006 | 1 | National Bureau of Statistics, Nigeria |
| Rwanda | Census | 2010 | 2 | United States Census Bureau, USAID |
| Senegal | Census microdata | 2002 | 2 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Sierra Leone | Census | 2004 | 2 | Statistics Sierra Leone |
| Somalia | MICS | 2006 | 1 | UNICEF |
| South Africa | Census | 2010 | 3 | United States Census Bureau, USAID |
| South Sudan | Census microdata | 2008 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Sudan | Census microdata | 2008 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Swaziland | DHS | 2007 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Togo | DHS | 2013 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Tunisia | Census | 2004 | 1 | National Statistics Institute, Tunisia |
| Uganda | DHS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| United Rep. of Tanzania | Census | 2010 | 2 | United States Census Bureau, USAID |
| Western Sahara | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Zambia | Census | 2010 | 3 | United States Census Bureau, USAID |
| Zimbabwe | DHS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |

Table 1. Data sources for African countries from which age and sex proportions were derived. Notes: AIS = Aids Indicator Survey; DHS = Demographic and Health Survey; MICS = Multiple Indicator Cluster Survey; MIS = Malaria Indicators Survey; the administrative unit level number in the 3rd column refers to the types of administrative divisions used for each country (with size of the administrative units decreasing from level 0, corresponding to the whole country, to level 3; <https://en.wikipedia.org/wiki/List_of_administrative_divisions_by_country>).

| **Country** | **Data type used** | **Year** | **Administrative Unit Level** | **Data source** |
| --- | --- | --- | --- | --- |
| Afghanistan | DHS | 2010 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Armenia | Census microdata | 2001 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Azerbaijan | DHS | 2006 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Bangladesh | DHS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Bhutan | Census | 2005 | 2 | National Statistics Bureau, Bhutan |
| Brunei Darussalam | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Cambodia | Census | 2010 | 2 | United States Census Bureau, USAID |
| China | Census | 2010 | 3 | National Bureau of Statistics of the People's Republic of China |
| Georgia | MICS | 2005 | 1 | MEASURE Demographic and Health Surveys, USAID |
| India | Census | 2010 | 1 | United States Census Bureau, USAID |
| Indonesia | Census | 2010 | 1 | United States Census Bureau, USAID |
| Iran | Census microdata | 2006 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Iraq | Census microdata | 1997 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Kazakhstan | MICS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Korea, Democratic People's Republic of | Census | 2010 | 1 | National Bureau of Statistics, North Korea |
| Kyrgyz Republic | Census microdata | 1999 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Laos | SIS | 2012 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Malaysia | Census microdata | 2000 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Mongolia | Census microdata | 2000 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Myanmar | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Nepal | DHS | 2011 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Pakistan | Census microdata | 1998 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Papua New Guinea | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Philippines | Census microdata | 2000 | 1 | Integrated Public Use Microdata Series, International (IPUMSI) |
| Russia | Census | 2010 | 1 | United States Census Bureau, USAID |
| Singapore | Census | 2010 | 2 | Statistics Singapore, Singapore |
| Solomon Islands | Census | 2009 | 1 | Solomon Islands National Statistics Office, Solomon Islands |
| South Korea | Census | 2010 | 1 | Statistics Korea, South Korea |
| Sri Lanka | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Tajikistan | DHS | 2012 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Thailand | Census | 2010 | 2 | United States Census Bureau, USAID |
| Timor-Leste | DHS | 2010 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Turkey | DHS | 1998 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Turkmenistan | UN | 2010 | 0 | World Population Prospects, United Nations Population Division |
| Uzbekistan | DHS | 1996 | 1 | MEASURE Demographic and Health Surveys, USAID |
| Vanuatu | Census | 2009 | 1 | Vanuatu National Statistics Office (VNSO), Vanuatu |
| Vietnam | Census microdata | 2009 | 2 | Integrated Public Use Microdata Series, International (IPUMSI) |

Table 2. Data sources for Asian countries from which age and sex proportions were derived. Notes: AIS = Aids Indicator Survey, DHS = Demographic and Health Survey, MICS = Multiple Indicator Cluster Survey, MIS = Malaria Indicators Survey; the administrative unit level number in the 3rd column refers to the types of administrative divisions used for each country (with size of the administrative units decreasing from level 0, corresponding to the whole country, to level 3; <https://en.wikipedia.org/wiki/List_of_administrative_divisions_by_country>).

**Data Preparation**

Depending on the format and structure of the raw tables containing the age and sex data for each given country, a specific processing technique was developed and applied in order to standardise all data tables across all countries. Table 3 presents an example of a standardised table, for Bhutan, in which each record represents a subnational administrative unit and the fields contain the corresponding proportionate values of people (both sexes) in each 5-year age group, males and females. There follows a summary of the main techniques employed to process raw table data, summarised by data source.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **NAME** | **ADM\_NAME** | **MPROP** | **FPROP** | **A0004** | **A0509** | **A1014** | **A1519** | **A2024** | **A2529** | **A3034** | **…** | **A65PL** |
| Bhutan | BUMTHANG | 54.3 | 45.7 | 8.5 | 9.9 | 12.4 | 12.1 | 12.1 | 9.2 | 6.4 | … | 6.2 |
| Bhutan | CHHUKHA | 56.9 | 43.1 | 9.4 | 10.5 | 10.7 | 10.6 | 13.4 | 11.7 | 8.2 | … | 2.6 |
| Bhutan | TSIRANG | 51.0 | 49.0 | 10.4 | 10.6 | 12.9 | 11.9 | 9.7 | 7.9 | 5.9 | … | 5.4 |
| Bhutan | DAGANA | 50.3 | 49.7 | 11.1 | 13.1 | 14.4 | 12.4 | 8.0 | 7.1 | 5.9 | … | 4.6 |
| Bhutan | GASA | 52.5 | 47.5 | 10.4 | 11.5 | 10.8 | 9.1 | 10.5 | 9.4 | 7.8 | … | 5.4 |
| Bhutan | SARPANG | 52.1 | 47.9 | 10.6 | 11.7 | 12.0 | 11.4 | 11.1 | 9.4 | 6.9 | … | 3.7 |
| Bhutan | HAA | 53.9 | 46.1 | 8.5 | 10.6 | 12.6 | 12.8 | 10.1 | 9.0 | 6.9 | … | 4.9 |
| Bhutan | LHUENTSE | 50.2 | 49.8 | 10.3 | 12.8 | 12.4 | 12.4 | 7.5 | 7.7 | 5.8 | … | 6.8 |
| Bhutan | MONGGAR | 50.4 | 49.6 | 10.3 | 11.8 | 12.8 | 12.8 | 9.0 | 7.8 | 6.2 | … | 5.6 |
| Bhutan | PARO | 53.0 | 47.0 | 8.4 | 10.0 | 11.4 | 11.9 | 12.6 | 9.6 | 7.1 | … | 5.3 |
| Bhutan | PEMAGATSHEL | 49.5 | 50.5 | 9.8 | 11.3 | 12.6 | 13.6 | 7.4 | 6.5 | 4.6 | … | 8.1 |
| Bhutan | PUNAKHA | 50.7 | 49.3 | 9.7 | 11.6 | 11.8 | 14.8 | 9.1 | 7.5 | 6.0 | … | 5.4 |
| Bhutan | SAMTSE | 52.1 | 47.9 | 10.0 | 11.3 | 12.5 | 10.8 | 10.6 | 8.7 | 6.8 | … | 4.9 |
| Bhutan | SAMDRUPJONGKHAR | 51.4 | 48.6 | 10.6 | 11.8 | 12.8 | 11.0 | 9.3 | 8.4 | 6.3 | … | 4.9 |
| Bhutan | ZHEMGANG | 50.9 | 49.1 | 9.7 | 11.6 | 14.0 | 13.6 | 8.8 | 6.7 | 5.4 | … | 6.5 |
| Bhutan | TRASHIYANGTSE | 49.9 | 50.1 | 10.9 | 12.4 | 13.0 | 12.3 | 8.1 | 6.8 | 5.4 | … | 5.3 |
| Bhutan | TRASHIGANG | 51.0 | 49.0 | 9.9 | 11.9 | 13.1 | 13.1 | 9.4 | 6.9 | 5.7 | … | 5.6 |
| Bhutan | THIMPHU | 54.2 | 45.8 | 9.6 | 9.6 | 10.9 | 11.6 | 15.3 | 11.4 | 7.9 | … | 3 |
| Bhutan | TRONGSA | 51.2 | 48.8 | 9.9 | 11.8 | 12.5 | 11.6 | 9.2 | 8.1 | 6.1 | … | 6.9 |
| Bhutan | WANGDUE | 51.7 | 48.3 | 10.1 | 11.5 | 12.3 | 11.3 | 10.1 | 8.1 | 6.6 | … | 5.7 |

Table 3. Example of a standardised country table containing the proportionate values of people (both sexes) in each 5-year age group, males and females in each administrative unit. (Note that the table is for illustrative purposes only, and hence only displays age groups up to 34 years of age and 65 plus).

***Processing National Census Data***

National census data is recorded and documented according to protocols determined by national governments; hence a wide range of different table data formats and structures needed to be processed and standardized. Microsoft Excel was used to manually restructure the raw data tables into a common format, structure and schema, similar to the one presented in Table 3.

**Processing IPUMSI data**

Unlike data derived from full censuses, IPUMSI data tables are already provided in a standard format and structure and thus a model was developed and applied in order to automate the processing of standardizing them. Raw country-level data are provided as comma-separated values (CSV) tables with accompanying spatial boundaries. Each CSV table is structured according to one row per person surveyed and the person’s age and sex is recorded in two fields in the same table. The raw table data were processed as follow:

* Convert each CSV table to a file geodatabase table
* Create new fields for each 5-year age group and sex class and populate with binary values, 0 or 1; a value of 1 in an age group field would indicate that the person's age fell in that range
* Use a summarise function to sum the new fields by administrative unit (note, a ‘count’ of all records by administrative unit also provides the total sampled population)
* Create and populate new fields for proportion of total population within each 5-year age group and sex class, by administrative unit
* Perform QA/QC to ensure that for each administrative unit a sum of people in all 5-year age groups equates to total population
* Perform QA/QC to ensure that for each administrative unit a sum of male and female equates to total population
* Join new table data to spatial boundaries based on administrative unit unique IDs
* Export the temporary join to a geodatabase feature class

The ArcGIS Toolbox containing the ArcGIS ModelBuilder models was used to perform all tasks described above and is distributed as part of the WorldPop-DepRatioAgeStruc-v1 code41 described in the Code availability subsection below.

**Processing Household Survey Data**

Household survey data are designed to be representative at both national and sub-national level, where the sub-national level usually corresponds to region or province, depending of the survey sample design and representativeness.

Information about age structure by region were derived from household survey data by using raw data files. Given that the format and structure of all household survey data are very similar, a data management process was developed to produce standardized outputs from the raw tables containing age and sex data. Household members’ files, which include information related to each household member's age, sex and administrative unit of residence, were accessed and downloaded from the relative web pages, namely, Measure Demographic and Health Surveys (DHS) program38 (which also includes MIS and AIS), UNICEF Multiple Indicator Cluster Survey (MICS)40 and Social Indicator Surveys (SIS)39. Data processing was mainly performed using the SPSS (IBM) statistical software42. Sampling weights were applied to the calculations in order to ensure representativeness, following the relative instructions given by the data providers’ documentation. The variable containing sampling weights was identified and made available in the appropriate SPSS format. Also, in order to account for the sampling strategy adopted by the survey, relevant variables for strata and primary sampling unit were defined in SPSS (using CSPLAN ANALYSIS). Following the DHS documentation and final report, age structures were calculated only on *de facto members*, which are generally defined as those household members who slept in the household the night before the interview. Moreover, variables in the datasets corresponding to members’ age, sex and administrative unit of residence were identified. After selecting only de facto members and weighting the data appropriately, cross tabulations were applied to calculate counts and proportions of population within each 5-year age group and sex class, by administrative unit. Finally, outputs of cross tabulations were exported into excel and reformatted in order to match the standard table schema described in the Data preparation subsection.

A template SPSS syntax file, showing the process for creating proportional age and sex structures from household survey data, is distributed as part of the WorldPop-DepRatioAgeStruc-v1 code41 described in the Code availability subsection below.

**Joining standardised data tables to spatial boundaries**

Once formatted to a standard table schema, subnational age and sex structure data were then joined on a country-by-country basis, using a GIS system, to their corresponding spatial boundaries; with the latter representing the administrative unit level at which the data were assembled. The ArcGIS Toolbox tool ‘Join Field’ was used for this purpose (…\data management tools.tbx\joins\join field). Note the exception for the IPUMSI data for this stage as processing that data was accomplished as part of the model workflow described in the Processing IPUMSI data subsection.

Spatial boundary datasets were obtained from a range of sources, including the GADM database43, the DHS Spatial Repository44 and multiple national statistical offices. Mismatches with subnational administrative units and topological inconsistences between national boundaries were manually corrected using a GIS system.

Supplementary Figures 1a and 1b (Africa and Asia, respectively) present the boundaries of the administrative units and each unit is coloured according to the proportion of the total population sampled and used to derive the 5-year age group and sex proportions. These country-level spatial datasets were then merged into two continental vector datasets (i.e. an African and Asian dataset) using the ArcGIS ‘Merge’ tool (toolboxes\system toolboxes\data management tools.tbx\General\Merge).

**Producing subnational dependency ratio datasets**

Using the two continental vector datasets described above, dependency ratios were calculated using a field calculation. The ArcGIS tool ‘Field Calculator’ was used to populate three new fields (i.e. CDR, YDR and ODR), listing subnational dependency ratio values calculated at the administrative unit level:

CDR = ((pc0\_14) + (pc65) / (pc15\_64)) x 100 (1)

YDR = ((pc0\_14) / (pc15\_64)) x 100 (2)

ODR = ((pc65) / (pc15\_64)) x 100 (3)

where CDR, YDR and ODR represent the combined, young and old dependency ratio, respectively, (pc0\_14) represents the proportion of the population aged 0 to 14, (pc65) represents the proportion of the population aged over 65, and (pc15\_64) represents the proportion of population aged 15 to 64.

All three subnational dependency ratio vector datasets were rasterized at a resolution of 30 arc seconds (approximately 1 km at the equator). Figure 2 illustrates the spatial distribution of the subnational YDR dependency ratios in Africa and Asia, respectively.

**Producing subnational 5-year age/sex group count datasets**

At this stage, to produce high resolution gridded 5-year age/sex group count datasets for 2010, the WorldPop continental gridded population count datasets for Africa and Asia (http://www.worldpop.org.uk/data/data\_sources/), with the total population for each country adjusted to match United Nations Population Division (UNPD) estimates for 201035, were multiplied by the comprehensive subnational proportional age and sex structures assembled.

The first step in this process was to convert the subnational administrative units, represented by polygon features in the two continental vector datasets described above, from vector format to raster grid format. Each polygon feature (representing a subnational administrative unit) is attributed with proportionate values for each 5-year age group and sex class (i.e. the proportion of the total population in each administrative unit belonging to each 5-year age group and sex class). The conversion process produced a stack of raster grids (one for each 5-year age group and sex class) and within each grid each pixel retained the proportionate value for that 5-year age group or sex class relating to the subnational administrative unit in which the pixel is located. Bearing in mind that subsequently these grids have to be overlaid on the WorldPop gridded continental population count datasets for Asia and Africa, in the interests of accurate grid cell calculations, the conversion process needs to ensure that the mesh of the grid is identical in terms of grid cell size (resolution) and grid cell alignment, to that of the WorldPop datasets. The conversion process uses the ArcGIS tool *Polygon to Raster*. A python script was created which called upon this tool and applied it, through iteration and whilst maintaining cell properties as described (using Geoprocessing Environment Settings), to all fields within the age/sex structure vector dataset containing the 5-year age group and male/female proportions. The ArcGIS Toolbox containing the Geoprocessing tool which calls the Python script is distributed as part of the WorldPop-DepRatioAgeStruc-v1 code41 described in the Code availability subsection below.

ArcGIS Model Builder was then used to automate the processing of the resulting stack of raster grids, (representing gridded subnational 5-year age group proportions). The model iterated through the stack such that each raster was sent to *Raster Calculator* tool along with the corresponding WorldPop continental gridded population count datasets (Africa or Asia) and both were utilised in a simple map algebra calculation:

AGC = AGP \* WPPC  *(4)*

(where AGC represents the resulting gridded population count dataset for a given 5-year age group, AGP is the gridded proportion dataset for the corresponding age group and WPPC is the WorldPop continental gridded population count dataset either for Africa or Asia).

This model produced a series of grids for Africa and Asia, each one providing estimates of population count for a specific age group (circa 2010) at the grid cell level. To disaggregate them by sex, two similar models, one for males and one for females, iterated through this stack to produce sex delineated population counts for each age group (circa 2010) at the grid cell level. These two models call upon the grids for subnational male and female proportions generated at an earlier stage as previously described.

Figure 3 represents an application of these datasets, providing gridded estimates of both young and working age population distribution (aged 0 to 14 and 15 to 65 age, respectively) for Africa and Asia.

The ArcGIS Toolbox containing the ArcGIS ModelBuilder models used to perform the steps described above is distributed as part of the WorldPop-DepRatioAgeStruc-v1 code41 described in the Code availability subsection below.

**Code Availability**

The WorldPop-DepRatioAgeStruc-v1 code41, used to produce the datasets described in this article, is publicly available through Figshare. It consists of (1) an ArcToolbox Geoprocessing Tool to pre-process the raw IPUMSI data, (2) an SPSS (IBM version 22) script to pre-process the DHS raw data and (3) an ArcToolbox Geoprocessing Tool to generate gridded age group structure and sex class proportion datasets and combine them with gridded population count datasets to produce the gridded age/sex structure count datasets. All of them are internally documented in order to both briefly explain their purpose and, when required, guide the user through their customization.

### Data Records

### All datasets described in this article, for the 87 countries listed in Table 1 and 2, are publicly and freely available through both the WorldPop Dataverse Repository (Data Citation 1, Data Citation 2, Data Citation 3, Data Citation 4) and the WorldPop website (<http://www.worldpop.org.uk/data/data_sources/>). While the datasets available through the WorldPop Dataverse Repository will be preserved in their published form while the ones available through the WorldPop website will be integrated with additional countries (Middle Eastern countries and Japan). Furthermore, additional gridded subnational dependency ratio datasets and high resolution gridded 5-year age/sex group count datasets for all countries located in Latin America and the Caribbean will be soon available through the WorldPop website.

### Gridded subnational dependency ratio datasets and high resolution gridded 5-year age/sex group count datasets for each country listed in Table 1 and 2 can be obtained by (i) downloading the corresponding datasets associated with the continent in which the country of interest is located and (ii) using the accompanying gridded ISO country code mask45 to extract them (Table 4).

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Description** | **Format/**  **Resolution** | **Dataverse DOI** |
| AFR\_2010\_SubNat\_DepRatio | Subnational combined (CDR), young (YDR) and old (ODR) dependency ratio datasets for 50 African countries | GeoTIFF/  3 arc seconds | 10.7910/DVN/S5JHQN |
| ASIA\_2010\_SubNat\_DepRatio | Subnational combined (CDR), young (YDR) and old (ODR) dependency ratio datasets for 37 Asian countries | GeoTIFF/  3 arc seconds | 10.7910/DVN/6TPPZ8 |
| AFR\_PPP\_AGE\_2010 | Female/male per pixel, in the age group from XX to YY, for 50 African countries in 2010; with total population for each country adjusted to match 2010 UN national estimates39 | GeoTIFF/  3 arc seconds | 10.7910/DVN/4MJN3G |
| ASIA\_PPP\_AGE\_2010 | Female/male per pixel, in the age group from XX to YY, for 37 Asian countries in 2010; with total population for each country adjusted to match 2010 UN national estimates39 | GeoTIFF/  3 arc seconds | 10.7910/DVN/GUSJUZ |
| WP\_DemographicData\_CountryMask | Numeric ISO-3166 country codes for all 87 African and Asian countries listed in Table1 an 2 | GeoTIFF/  3 arc seconds | 10.7910/DVN/S5JHQN  10.7910/DVN/6TPPZ8  10.7910/DVN/4MJN3G  10.7910/DVN/GUSJUZ |

Table 4. Name, description (from XX to YY represents all 5-year age groups from 00 up to a final group aged 65 and over), format, resolution and DOI of all datasets available for Africa and Asia (including the gridded ISO country code mask45 needed to extract country datasets from the continental datasets).

### Technical Validation

All data collected, assembled and used were (i) already validated by the corresponding data collector, owner and/or distributor, and (ii) further checked, in the framework of this project, to ensure that they represent true trait variation by inspecting the proportions in each age class and sex group (making sure they were within reasonable ranges on a country-by-country basis). Then, both the gridded subnational dependency ratio datasets and high resolution gridded 5-year age/sex group count datasets were produced by solely processing the input data and thus the outputs were simply verified by inspecting them for abnormal values and nonsensical results. For the high resolution gridded 5-year age/sex group count datasets, both for Africa and Asia, this was done by summing all of them into a single dataset (depicting the total numbers of people for all age groups and both sexes at the grid cell level) and then subtracting it from the corresponding WorldPop continental gridded population count dataset to make sure that country totals matched the UNPD estimates for 2010. Subnational dependency ratios were also evaluated by examining their spatial distribution and by comparing them against existing UN country-level estimates.

**Comparing sub-national dependency ratios against national estimates**

The distributions of the YDR at sub-national scales for Africa and Asia were plotted against World Bank national estimates, for each country’s correspondent year. Figure 4 gives an indication of the size of the sub-national variation in the YDR shown in the subnational level datasets that is masked when averaging at national levels.

The variation observed is related also to the administrative unit level of the input data, with those countries for which the most spatially detailed YDR were available (i.e. Tanzania, South Africa and China) showing the largest differences between minimum and maximum estimates. Nevertheless, even for those countries where subnational data were only available at administrative unit level 1, sub-national differences from the World Bank national level estimates are evident. Additionally, national estimates are often highly influenced by areas with high density of population whose rates greatly influence the averages, and this is reflected by the World Bank estimates falling outside the interquartile range of the boxplots in many cases (Figure 4).

**Sensitivity tests on differences over time in population structures**

All available sources that were as close as possible to 2010 at the time of data processing were considered in this work. However, for some countries only relatively outdated data were available. To test the assumption of no major changes in age structures between years, the absolute change (difference) in age group specific percentage points was tested, in countries where two or more comparable time points were available from household surveys,. Two countries in Asia and one country in Sub Saharan Africa were selected, namely: Nepal (DHS Surveys 2011-2006 & 2011-2001), Kazakhstan (MICS Surveys 2011-2006) and Guinea (DHS Surveys 2011-2005 & 2011-1999). Absolute change (difference) in percentage points was calculated for each survey region and total country (Table 5).

Also, for the same countries and all different time points, population pyramids were constructed and structures between years were compared (Supplementary Figures 2a-4c). Population pyramids constructed for this test follow the methodology described in the Methods section, therefore they reflect the input data utilised to construct the datasets described in this article. In order to be consistent with the method applied here, country level proportions of males and females were applied for all age groups. Results from our tests suggest that overall, differences are relatively small. Through looking at country totals and comparing pyramids, the highest differences in percentage points are shown for younger age groups, and this pattern is consistent among all tested countries. Furthermore, country level tests for Kazakhstan also show that some older age groups (60-64 and 65-69) present larger differences in percentage points, compared to the other age groups. At subnational level, the most consistent differences are shown in Central, Far-western regions in Nepal when comparing the 2001 to the 2011 survey, and in Mangystau oblysy (Kazakhstan 2011-2006) for age group 0-5, when compared to other regions and age classes.

### Usage Notes

The gridded subnational dependency ratio datasets and high resolution gridded 5-year age/sex structure datasets described in this article can be used to support a range of applications, from planning interventions to designing strategies and deriving health/development metrics, and to predict response variables that are intrinsically dependent on dependency ratios and age/sex structures.

Ongoing work involves the integration of these datasets with information related to health and development for planning interventions. This can include, for example, integration with information related to vaccination rates of children under the age of five26, access to antenatal, delivery and postnatal care for women of reproductive age46, and disease prevalence and burden47,48.

These datasets can also be used as a base for measuring demographic progress in relation to a set of topics like, ageing, population growth, dynamics and projections5,6,8,15,49,50. Furthermore, they can be combined with economic data to assess the association between the population pyramid and the economic development in a given country, as well as to study the effects of population structure on savings and growth51. Finally, age and sex compositions can be used to determine the demographic events that will occur and their impact on the types of facilities and services (e.g., schools and maternity services) needed by the population.

Although the subnational dependency ratio and age/sex structure datasets presented here represent the best available datasets at this level of spatial detail/resolution, it is important to highlight that limitations still exist. Indeed, all of the census data, census microdata and survey-based data used here are subject to various sources of error and bias.

Surveys usually do not sample from certain groups and places (e.g. indigenous groups, informal settlements, places experiencing civil unrest and refugee camps) either because of political biases, missing sampling frames, or security issues. Therefore, these groups are not covered in surveys, and in some cases are also not included in national censuses. Household survey data are in most cases sampled using complex survey design procedures52 and estimates derived from household survey datasets are usually representative at administrative level 1 or provinces, or a combination of those. IPUMS-International microdata are samples from population censuses from National Statistical Offices taken around the world since 1960. Sample data are derived by using different sampling designs (<https://international.ipums.org/international/variance_estimation.shtml>). Furthermore, they cover only a fraction of the population (approximately 5%) and, as for every sample, they are subject to sampling errors 37.

Uncertainties also arise over comparisons being made between primarily full census-based estimates of dependency ratios and age/sex structures and those based on IPUMSI microdata and household survey data. Differences between the way these were measured contributes to uncertainties in comparisons, though strong correlations between household survey-derived age structures, and those derived from census data suggest that such differences may be small53. For some countries, the input data used in this project were also relatively outdated and/or coarse - whilst the precision with which heterogeneities in dependency ratios and age/sex structures are mapped is improved over simple country-level estimates, the datasets presented here are still limited through e.g. only one set of values for Libya and thousands for China. Moreover, like most other population characteristics reported at the administrate unit level, the dependency ratios and age/sex structures are also subject to the modifiable areal unit problem54.

Population data that were sex disaggregated by age group at subnational administrative unit level were only available for a minority of countries. In most cases, the subnational level data were disaggregated only by age group with a single two-way sex split. Therefore in order to maintain the comparability across countries, the subnational level sex split was simply applied across all age groups within the corresponding subnational administrative unit.

**Acknowledgements**

This work is supported by funding from the Bill & Melinda Gates Foundation (OPP1134076). CP is supported by funding from the Bill & Melinda Gates Foundation (OPP1106427). GMH and AS are supported by funding from the Bill & Melinda Gates Foundation (OPP1106427, 1032350). AEG is supported by funding from Google (OICB150153) and the Bill & Melinda Gates Foundation (OPP1134076). CL is supported by funding from the Belgian Science Policy (SR/00/304). AJT is supported by funding from NIH/NIAID (U19AI089674) and acknowledges the support of the RAPIDD program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health. AJT is also supported by the Bill & Melinda Gates Foundation (OPP1106427, 1032350, OPP1134076, OPP1094793), the Clinton Health Access Initiative as well as a Wellcome Trust Sustaining Health Grant (106866/Z/15/Z). This work forms part of the WorldPop Project ([www.worldpop.org.uk](http://www.worldpop.org.uk/)) and Flowminder Foundation ([www.flowminder.org](http://www.flowminder.org/)).

**Contributions**

C.P, G.H.M. and A.S. drafted the manuscript. C.P., G.M.H., A.S., A.E.G., T.J.B., C.T.L . and D.K. acquired and assembled the raw data. G.M.H., A.S. and D.K. produced the WorldPop gridded continental population count datasets for 2010. G.M.H. and D.K. produced the final datasets. C.P., G.H.M. and D.K. performed the technical validation of the raw data and final datasets. G.H.M, A.S. and C.T.L. produced the gridded ISO country code datasets accompanying the final datasets. C.P. and G.M.H. prepared the tables and figures. A.J.T. aided drafting the manuscript. C.P., G.M.H., C.T.L. and A.J.T. conceived the study. All authors read and approved the final version of the manuscript.

### Competing interests

The authors declare that they have no competing financial interests that might have influenced the presentation of the datasets and the description of the methods used to produce them.

### Data Citations

1. Pezzulo, C., Hornby, G.M., Sorichetta, A., Gaughan, A.E., Linard, C., Bird, T.J., Kerr, D., Lloyd, C.T. & Tatem, A.J. *Harvard Dataverse*. http://dx.doi.org/10.7910/DVN/S5JHQN (2017).
2. Pezzulo, C., Hornby, G.M., Sorichetta, A., Gaughan, A.E., Linard, C., Bird, T.J., Kerr, D., Lloyd, C.T. & Tatem, A.J. *Harvard Dataverse*. http://dx.doi.org/10.7910/DVN/6TPPZ8 (2017).
3. Pezzulo, C., Hornby, G.M., Sorichetta, A., Gaughan, A.E., Linard, C., Bird, T.J., Kerr, D., Lloyd, C.T. & Tatem, A.J. *Harvard Dataverse*. http://dx.doi.org/10.7910/DVN/4MJN3G (2017).
4. Pezzulo, C., Hornby, G.M., Sorichetta, A., Gaughan, A.E., Linard, C., Bird, T.J., Kerr, D., Lloyd, C.T. & Tatem, A.J. *Harvard Dataverse*. http://dx.doi.org/10.7910/DVN/GUSJUZ (2017).

References

1 United Nations. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. *Department of Economic and Social Affairs, Population Division (UNDESA)* (2015).

2 United Nations. Population Dynamics in the Post-2015 Development Agenda: Report of the Global Thematic Consultation on Population Dynamics. *United Nations Population Fund* (2013).

3 Herrmann, M. Factoring population dynamics into sustainable development. *Development Co-operation Report 2012: Lessons in Linking Sustainability and Development (OECD)*, doi:10.1787/dcr-2012-11-en (2012).

4 Das Gupta M, Bongaarts J & Cleland J. *Population, Poverty, and Sustainable Development: A Review of the Evidence.* (World Bank, 2011).

5 Bongaarts, J. in *Population Matters: Demographic Change, Economic Growth, and Poverty in the Developing World* (Oxford University Press, 2001).

6 Bloom, D. E. & Canning, D. Demographics and Development Policy. *Development Outreach* **13**, 77-81 (2011).

7 Cohen, J. E. Human Population: The Next Half Century. *Science* **302**, 1172-1175, doi:10.1126/science.1088665 (2003).

8 United Nations. World Population Policies 2013. *Department of Economic and Social Affairs, Population Division (UNDESA)* (2013).

9 Shryock, H. S. & Siegel, J. S. *The methods and materials of demography.* (US Dept. of Commerce, Bureau of the Census, 1973).

10 United Nations. Definition of the indicators of population ageing in world population ageing: 1950-2050 - Annex I. *United Nations Publications* (2002).

11 Bloom, D. E., Canning, D. & Sevilla, J. The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change. *Population Matters Series.* (2003).

12 Bloom, D. E., Canning, D., Fink, G. & Finlay, J. E. Microeconomic Foundations of the Demographic Dividend. *PGDA Working Paper* **93** (2012).

13 Williamson, J. in *Population Matters: Demographic Change, Economic Growth and Poverty in the Developing World.* (eds N. Birdsall, A. Kelley, & S. Sinding) (Oxford University Press, 2001).

14 Schultz, T. P. Demographic Determinants of Savings: Estimating and Interpreting the Aggregate Association in Asia. *IZA Discussion Paper No. 1479; Yale University Economic Growth Center Discussion Paper Series No. 901.* (2005).

15 Leahy, E., Engelman, R., Vogel, C. G., Haddock, S. & Preston, T. *The Shape of Things to Come: Why Age Structure Matters to a Safer, More Equitable World.* (2007).

16 United Nations. Demographic Indicators: Dependency Ratio. *Department of Economic and Social Affairs, Population Division (UNDESA)* (2013).

17 United Nations. Human Development Report 2013 - The Rise of the South: Human Progress in a Diverse World. *United Nations Development Programme (UNDP)* (2013).

18 Van der Ven, R. & Smits, J. *The demographic window of opportunity: age structure and sub-national economic growth in developing countries.* (Nijmegen Center for Economics (NiCE) - Institute for Management Research - Radboud University Nijmegen, 2011).

19 Eastwood, R. & Lipton, M. Demographic transition in sub-Saharan Africa: How big will the economic dividend be? *Population Studies* **65**, 9-35, doi:10.1080/00324728.2010.547946 (2011).

20 Madhavan, S., Schatz, E. & Clark, B. Effect of HIV/AIDS-related mortality on household dependency ratios in rural South Africa, 2000-2005. *Population studies* **63**, 37-51, doi:10.1080/00324720802592784 (2009).

21 Fayissa, B. & Gutema, P. *Dependency Ratio and the Economic Growth Puzzle in Sub-Saharan Africa.* (Murfreesboro, Tenn. : Middle Tennessee State Univ., Dep. of Economics and Finance, 2010).

22 Hadley, C., Belachew, T., Lindstrom, D. & Tessema, F. The shape of things to come? Household dependency ratio and adolescent nutritional status in rural and urban Ethiopia. *American journal of physical anthropology* **144**, 643-652, doi:10.1002/ajpa.21463 (2011).

23 File, T. & Kominski, R. *Dependency Ratios in the United States: A State and Metropolitan Area Analysis. Data from the 2009 American Community Survey.* (Social, Economic, and Household Statistics Division (SEHSD) - U.S. Census Bureau, 2012).

24 Tatem, A. J. *et al.* Millennium development health metrics: where do Africa’s children and women of childbearing age live? *Population Health Metrics* **11** (2013).

25 Rossouw, S. & Webber, D. Millennium Development Goals: A Concern Over Sub-National Variations. *International Business & Economics Research Journal* **10** (2011).

26 Alegana, V. A. *et al.* Fine resolution mapping of population age-structures for health and development applications. *Journal of The Royal Society Interface* **12**, doi:10.1098/rsif.2015.0073 (2015).

27 Tatem, A. J. *et al.* Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation. *Population Health Metrics* **10**, 1-14, doi:10.1186/1478-7954-10-8 (2012).

28 Bigman, D. & Srinivasan, P. V. Geographical targeting of poverty alleviation programs: methodology and applications in rural India. *Journal of Policy Modeling* **24**, 237-255, doi:10.1016/s0161-8938(02)00108-4 (2002).

29 Elbers, C., Fujii, T., Lanjouw, P., Oezler, B. & Yin, W. Poverty alleviation through geographic targeting: How much does disaggregation help? *Journal of Development Economics* **83**, 198-213, doi:10.1016/j.jdeveco.2006.02.001 (2007).

30 United Nations. Sub-National Estimates of Human Capital Indicators: Localizing Investments for the Demographic Dividend. *United Nations Population Fund* (2016).

31 Bassett, T. J. & Winter-Nelson, A. in *The Atlas of World Hunger* Ch. 28, (University of Chicago Press, 2007).

32 Admassie, A. Explaining the High Incidence of Child Labour in Sub–Saharan Africa. *African Development Review* **14**, 251-275, doi:10.1111/1467-8268.00054 (2002).

33 Vijayakuma, S. An Empirical Study on the Nexus of Poverty, GDP Growth, Dependency Ratio and Employment in Developing Countries. *Journal of Competitiveness* **5**, 67-82, doi:10.7441/joc.2013.02.05 (2013).

34 Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* **4**, doi:10.1038/sdata.2017.4 (2017).

35 United Nations. World Population Prospects: The 2015 Revision. Population indicators. *United Nations Department of Economic and Social Affairs Population Division (UNDESA)* (2016).

36 United Nations Development Programme (UNDP). *World Urbanization Prospects: The 2014 Revision. CD-ROM Edition.*, <http://esa.un.org/unpd/wup/CD-ROM/> (2014).

37 Minnesota Population Center. *Integrated Public Use Microdata Series, International: Version 6.4 [Machine-readable database].* <https://international.ipums.org/international/citation.shtml> (University of Minnesota, 2015).

38 ICF International. *Demographic and Health Surveys (Various) [Datasets].* (Rockville, Maryland: Icf International [Distributor], 2000-2015).

39 Ministry of Health, Lao Statistics Bureau & UNICEF. *Lao PDR - Social Indicator Survey 2011-2012 [Dataset].* <http://microdata.worldbank.org/> (2012).

40 (UNICEF), U. N. I. C. s. F. *Multiple Indicator Cluster Survey (MICS) [Various Datasets].* <http://www.childinfo.org/mics4_surveys.html> (United Nations, 2000-2015).

41 Pezzulo, C. *et al.* *Source code for: Sub-national mapping of population pyramids and dependency ratios in Africa and Asia.*, doi:10.6084/m9.figshare.4596187.v1 (2017).

42 IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp. (Released 2013).

43 GADM. *Database of Global Administrative Areas.* <http://www.gadm.org/> (2015).

44 ICF International. *Spatial Data Repository, the DHS Program (Various) [Datasets].* (Rockville, Maryland: Icf International [Distributor], 2000-2015).

45 Christopher T. Lloyd, Sorichetta, A. & Tatem, A. J. High resolution global gridded data for use in population studies. *Scientific Data* **4**, doi:10.1038/sdata.2017.1 (2017).

46 Tatem, A. J. *et al.* Mapping for maternal and newborn health: the distributions of women of childbearing age, pregnancies and births. *International Journal of Health Geographics* **13**, 2, doi:10.1186/1476-072x-13-2 (2014).

47 Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* **380**, 2095-2128 (2012).

48 Wang, H. *et al.* Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* **380**, 2071-2094, doi:10.1016/S0140-6736(12)61719-X (2012).

49 United Nations. World Population Ageing, 1950-2050. *Department of Economic and Social Affairs, Population Division (UNDESA)* (2002).

50 Madsen, E., Daumerie, B. & Hardee, K. *The effects of age structure on development. Policy and issue brief.* (Population Action International, 2010).

51 Cruz, M. & Amer, A., S. On the Impact of Demographic Change on Growth, Savings, and Poverty. *World Bank* (2016).

52 ICF International. Demographic and Health Survey Sampling and Household Listing Manual. *MEASURE DHS, Rockville, Maryland: ICF International* (2012).

53 Tatem, A. J. *et al.* Millennium development health metrics: where do Africa’s children and women of child bearing age live? *Popul Health Metrics* **11**, doi:10.1186/1478-7954-11-11 (2013).

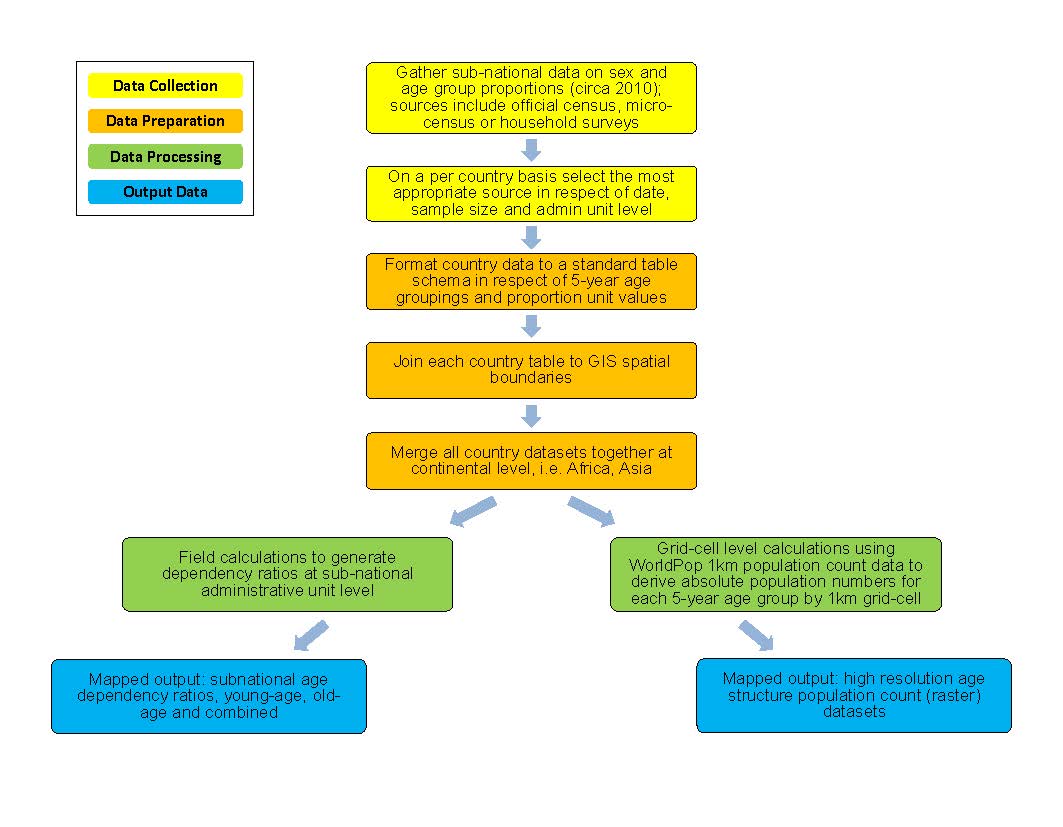
54 Openshaw, S. *The Modifiable Areal Unit Problem.* (Geo Books, 1984).

Figure 1: Schematic overview of the data processing method adopted to generate the WorldPop gridded subnational dependency ratio datasets and high resolution gridded 5-year age/sex group count datasets.

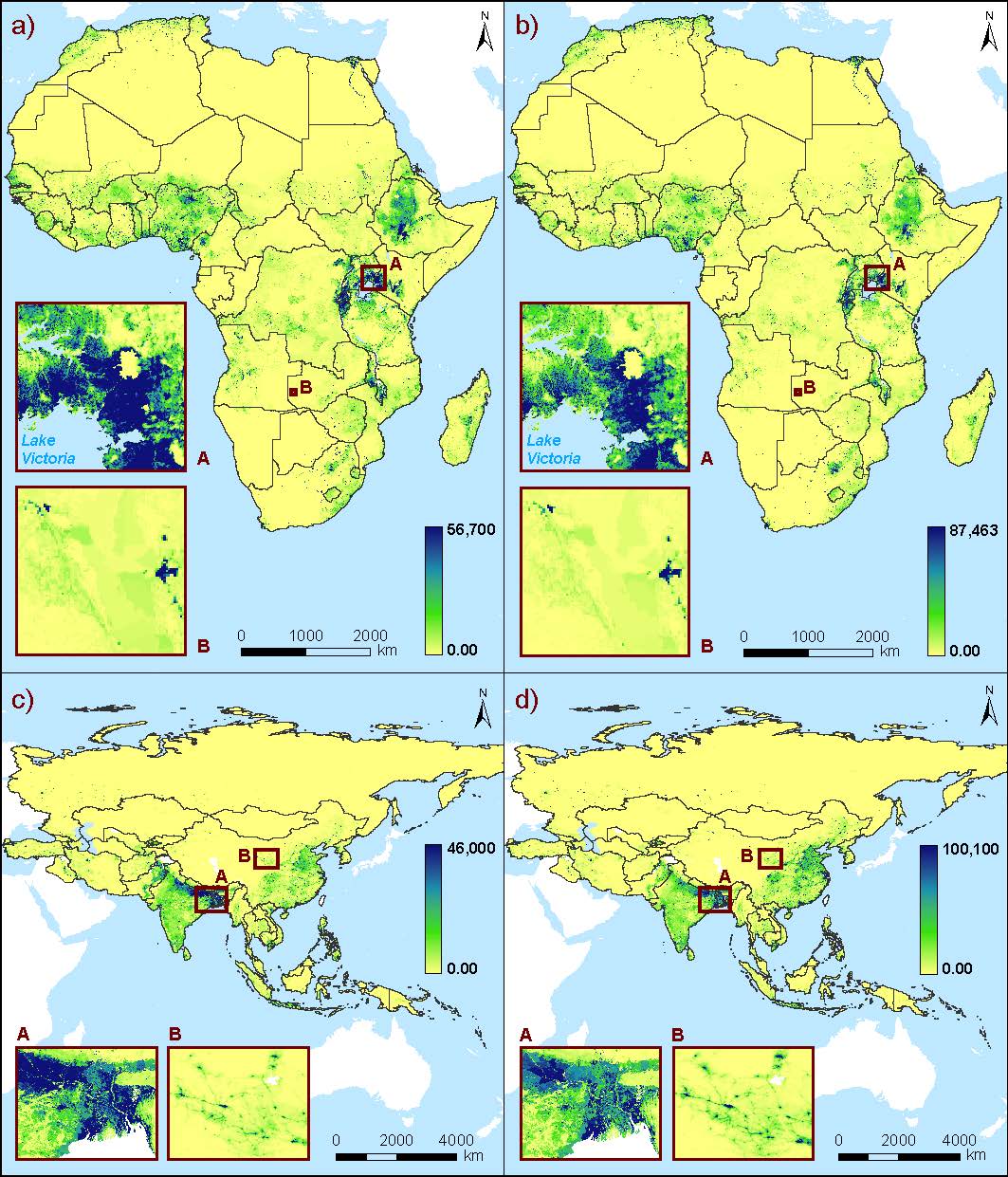
Figure 2: Subnational young age dependency ratio (YDR) datasets, circa 2010. Panel a) shows dataset for Asia and panel b) shows dataset for Africa. Note differing colour scales used between the two maps to highlight variations.

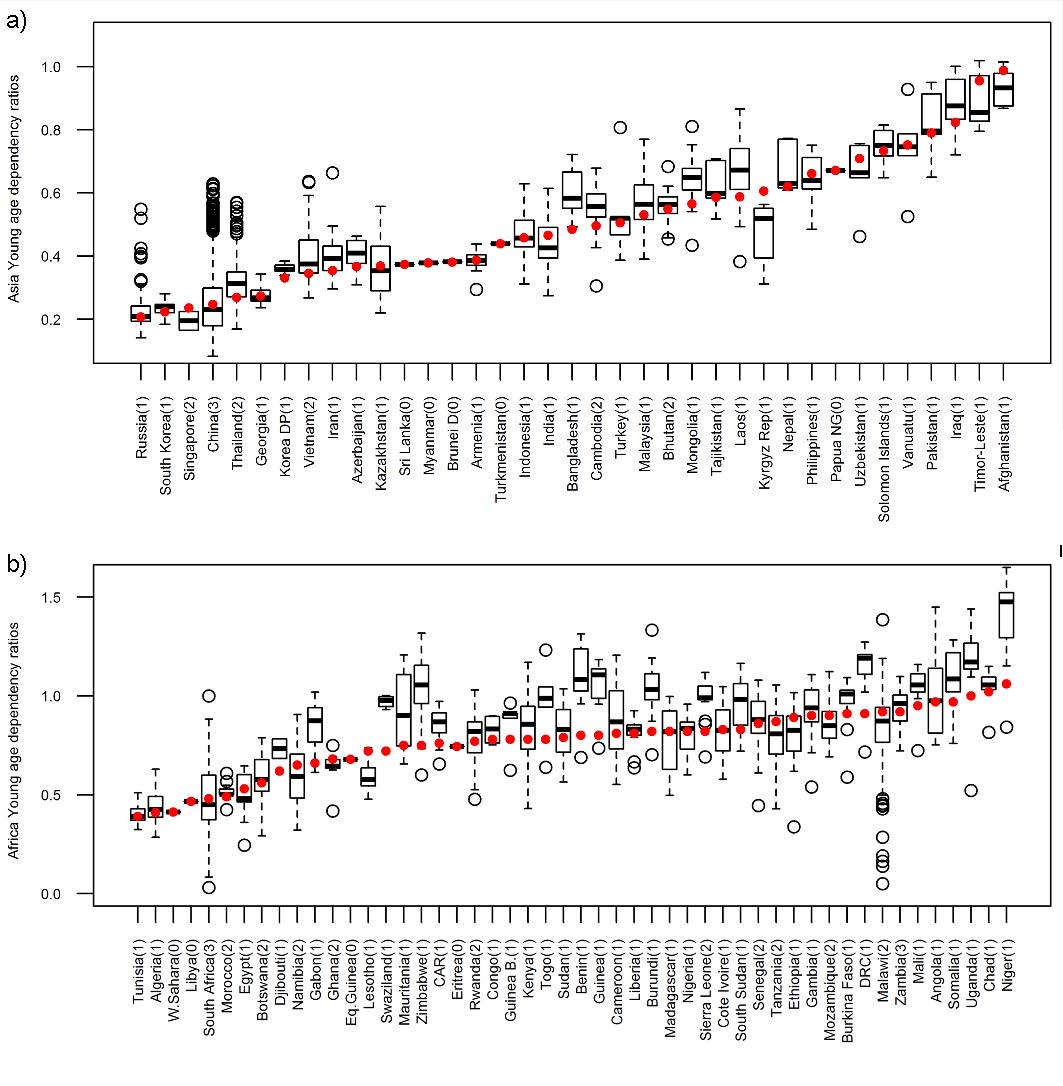
Figure 3. Circa 2010 High-resolution gridded population distribution presenting age structures for mainland Africa and Madagascar and the Asian region. Estimates of young age population (0 to 14yrs) are in panels (a) and (c); estimates of working age population (15 to 65) are in panels (b) and (d). The grid cell resolution is 30 arc seconds (approximately 1km at the equator) and coordinate reference system is GCS WGS 1984.

Figure 4. Differences between national and subnational young-age dependency ratios (YDR) for Asian and African countries. Panel a) shows Asian countries and panel b) shows African countries. For each country, the year of World Bank national estimates (red dots) correspond to the year of our data. The solid bold center line of the boxplot shows the median values, the box width represents the interquartile range, and the whiskers extend to 1.5 times the interquartile range from the box (values further away than this are shown as open circles). The administrative unit level of the sub-national data used here is shown in brackets next to the country name on the x-axis.









**Supplementary Figures**

**Table of Contents**

Supplementary Figure 1a…………………………………………………………………………….…………………………….………2

Supplementary Figure 1b……………………………………………………………………………………………………….…….……3

Supplementary Figure 2a……………………………………………………………………….…………………………….……..…….4

Supplementary Figure 2b……………………………………………………………………….…………………………….……..…….5

Supplementary Figure 2c……………………………………………………………………….…………………………….……..…….6

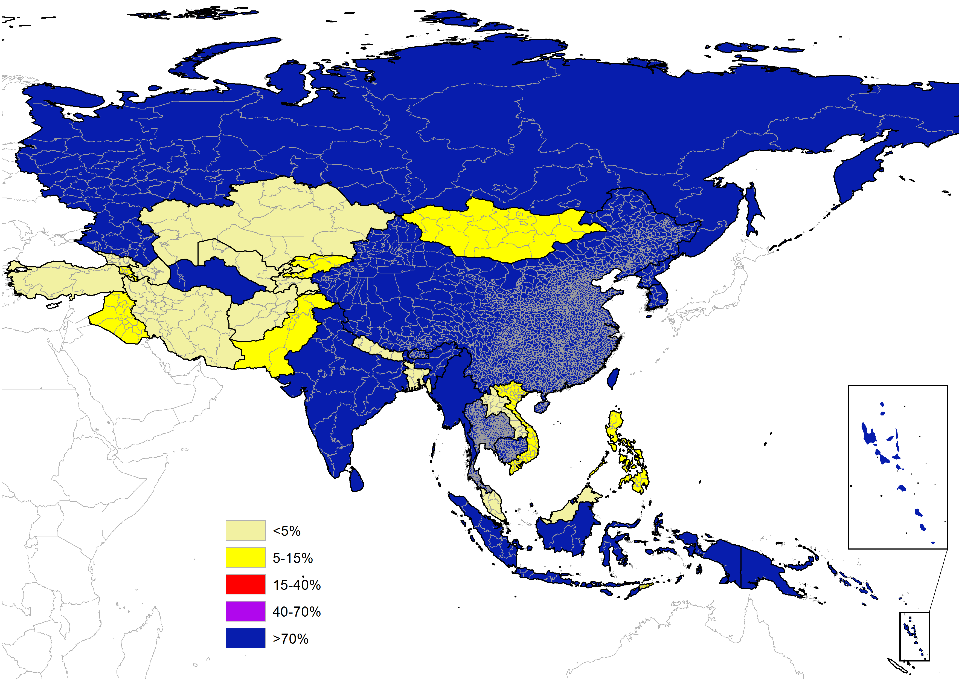
Supplementary Figure 3a……………………………………………………………………………………………………….…….……7

Supplementary Figure 3b……………………………………………………………………………………………………….…….……8

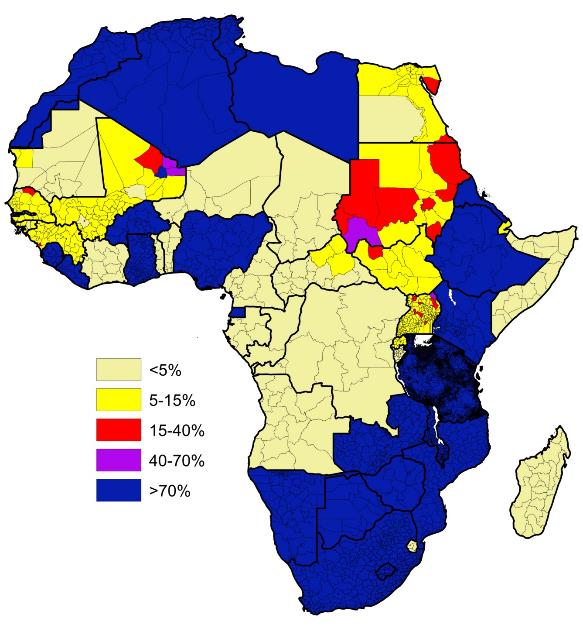
Supplementary Figure 4a…….……………………………………………………………………….……………………………………9

Supplementary Figure 4b…….……………………………………………………………………….………………………………….10

Supplementary Figure 4c…….……………………………………………………………………….………………………………….11



Supplementary Figure 1a. *Sample size of the sub-national age and sex proportion data used, displayed as a percentage of the estimated total sub-national unit population at the time of survey, Asia.*

**

Supplementary Figure 1b. *Sample size of the sub-national age and sex proportion data used, displayed as a percentage of the estimated total sub-national unit population at the time of survey, Africa.*



Supplementary Figure 2a. *Population pyramid, Nepal DHS, 2011. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 2b. *Population pyramid, Nepal DHS, 2006. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 2c. *Population pyramid, Nepal DHS, 2001. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 3a. *Population pyramid, Kazakhstan MICS, 2010. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 3b. *Population pyramid, Kazakhstan MICS, 2006. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 4a. *Population pyramid, Guinea DHS, 2012. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 4b. *Population pyramid, Guinea DHS, 2005. Population pyramids follow methodology as described in the manuscript (Methods section).*



Supplementary Figure 4c. *Population pyramid, Guinea DHS, 1999. Population pyramids follow methodology as described in the manuscript (Methods section).*

| **Country/Region** | **Age groups** | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **<5** | **5-9** | **10-14** | **15-19** | **20-24** | **25-29** | **30-34** | **35-39** | **40-44** | **45-49** | **50-54** | **55-59** | **60-64** | **65-69** | **70-74** | **75-79** | **80+** |
| **Nepal DHS (2011-2006)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Eastern | -2.0 | -1.4 | 0.5 | -0.6 | -0.2 | 0.4 | 0.6 | 0.3 | -0.1 | -0.4 | 0.9 | 0.9 | 0.0 | 0.5 | 0.0 | 0.5 | 0.2 |
| Central | -2.0 | -1.4 | 0.2 | 0.3 | 0.1 | -0.1 | 0.6 | 0.7 | 0.2 | -0.5 | 0.9 | 0.5 | -0.3 | 0.3 | 0.4 | 0.1 | 0.2 |
| Western | -2.6 | -2.5 | 0.0 | 0.7 | -0.4 | -0.2 | 0.1 | 0.6 | 0.4 | 0.5 | 0.9 | 0.9 | 0.7 | 0.5 | -0.1 | 0.2 | 0.3 |
| Mid-western | -0.3 | -1.8 | -0.5 | -0.2 | 0.7 | 0.7 | -0.1 | 0.1 | -0.7 | -0.2 | 0.8 | 0.8 | -0.2 | 0.4 | 0.1 | 0.3 | 0.0 |
| Far-western | -1.2 | -0.5 | 0.6 | -1.3 | 0.1 | -0.6 | -1.2 | 0.2 | 0.1 | -0.2 | 0.6 | 0.3 | 1.3 | 0.8 | 0.5 | 0.1 | 0.3 |
| **Total** | -1.9 | -1.6 | 0.2 | -0.1 | 0.0 | 0.0 | 0.2 | 0.5 | 0.1 | -0.2 | 0.9 | 0.7 | 0.2 | 0.5 | 0.2 | 0.2 | 0.2 |
| **Nepal DHS (2011-2001)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Eastern | -3.06 | -1.93 | -0.50 | -0.19 | 1.06 | 0.63 | -0.46 | 0.13 | -0.15 | -0.10 | 1.42 | 0.95 | 0.60 | 0.14 | 0.33 | 0.60 | 0.57 |
| Central | -5.35 | -3.47 | 1.13 | 1.24 | 0.22 | -0.02 | 0.60 | 0.87 | 1.03 | 0.63 | 0.99 | 0.71 | 0.29 | 0.19 | 0.36 | 0.07 | 0.56 |
| Western | -3.87 | -3.85 | 0.16 | 0.57 | 0.14 | -0.27 | 0.38 | 1.04 | 0.53 | 0.39 | 1.04 | 0.93 | 0.87 | 0.42 | 0.50 | 0.51 | 0.56 |
| Mid-western | -2.90 | -1.67 | -0.38 | 0.30 | 0.36 | 0.63 | -0.26 | 0.96 | 0.32 | -0.04 | 0.36 | 0.78 | 0.34 | 0.34 | 0.51 | 0.44 | -0.09 |
| Far-western | -4.33 | -1.49 | 0.28 | 1.12 | 0.60 | 0.64 | -0.62 | 0.53 | 0.38 | -0.01 | 0.70 | 0.28 | 0.53 | 0.46 | 0.44 | 0.18 | 0.35 |
| **Total** | -4.09 | -2.74 | 0.23 | 0.61 | 0.47 | 0.25 | 0.06 | 0.71 | 0.49 | 0.26 | 0.99 | 0.78 | 0.52 | 0.27 | 0.41 | 0.35 | 0.47 |
| **Kazakhstan (2011-2006)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Akmola oblysy | 0.13 | -0.69 | -1.81 | -1.58 | 0.48 | 0.61 | 0.19 | -0.78 | -0.76 | -0.51 | 2.44 | 0.39 | 2.07 | -1.36 | 1.73 | -0.72 | 0.19 |
| Aktobe oblysy | 2.90 | -0.31 | -1.91 | -2.11 | -0.08 | -0.23 | 1.06 | -0.83 | 0.07 | -0.10 | 0.67 | 0.92 | 1.12 | -2.21 | 1.41 | -0.57 | 0.21 |
| Almaty oblysy | 0.33 | 0.47 | -2.13 | -0.29 | 1.44 | -0.41 | -0.61 | -0.93 | 0.11 | 0.67 | 1.04 | -0.05 | 1.00 | -1.05 | 0.33 | -0.31 | 0.39 |
| Almati city | -2.82 | 0.99 | -2.00 | -1.99 | 2.35 | 1.08 | 0.88 | 2.55 | -0.69 | -0.17 | -1.89 | -0.49 | 2.00 | -1.48 | 0.83 | 0.36 | 0.49 |
| Astana city | 2.06 | 0.75 | -2.11 | -0.83 | 1.60 | 0.72 | -0.04 | -0.11 | -1.26 | -0.28 | 1.57 | -0.14 | 0.68 | -1.61 | 0.20 | -0.72 | -0.49 |
| Atyrau oblysy | 3.27 | 0.30 | -2.40 | -4.20 | 0.34 | 0.31 | 0.67 | -0.66 | -0.42 | 0.12 | 2.00 | 1.63 | 1.13 | -1.02 | -0.49 | -0.55 | -0.01 |
| Shygys-Kazakhstan oblysy | 2.36 | 1.37 | -0.77 | -3.18 | -1.22 | -0.37 | 1.24 | 0.14 | -0.22 | -0.61 | -0.63 | 1.38 | 1.79 | -1.98 | 0.83 | -0.64 | 0.51 |
| Zhambyl oblysy | 1.59 | 1.53 | -2.83 | 0.04 | 0.43 | -1.05 | -0.01 | -1.17 | -0.32 | 0.46 | 0.85 | -0.39 | 2.03 | -0.84 | 0.61 | -0.80 | -0.11 |
| Batys-Kazakhstan oblysy | 2.74 | 1.51 | -2.75 | -3.48 | -0.86 | 0.62 | 0.79 | -0.21 | -2.29 | -0.21 | 0.74 | 1.05 | 1.82 | -1.93 | 1.37 | 0.54 | 0.58 |
| Karagandy oblysy | 2.91 | 1.60 | -0.95 | -3.09 | -0.14 | 0.43 | 0.91 | 0.65 | -1.69 | -0.43 | -0.28 | 0.25 | 0.92 | -2.63 | 1.58 | 0.04 | -0.06 |
| Kostanai oblysy | 0.57 | 0.45 | -2.08 | -3.03 | 1.32 | 0.35 | -1.09 | -0.44 | -0.01 | 0.76 | 2.90 | -0.83 | 0.83 | -2.31 | 2.22 | -0.32 | 0.72 |
| Kyzylorda oblysy | 2.90 | 1.12 | -2.46 | -2.73 | -1.51 | -0.60 | 1.27 | 0.21 | -0.53 | 0.20 | 1.29 | -0.53 | 1.01 | -0.86 | 1.18 | 0.01 | 0.04 |
| Mangystau oblysy | 5.29 | 1.38 | -1.82 | -1.01 | -0.51 | -0.21 | 1.40 | -0.62 | -1.80 | -0.32 | -2.13 | 0.60 | 0.46 | -1.06 | 0.59 | -0.22 | -0.04 |
| Pavlodar oblysy | 1.38 | -0.21 | -2.28 | -2.98 | 0.43 | 2.07 | 1.38 | -0.98 | -1.95 | 0.58 | 1.60 | 0.23 | 2.88 | -2.55 | 1.00 | -0.33 | -0.26 |
| Soltustik-Kazakhstan oblysy | -0.08 | 0.97 | -1.23 | -3.21 | -0.23 | -0.46 | 0.24 | 1.71 | -1.20 | -0.74 | 0.84 | 1.87 | 2.03 | -2.40 | 1.70 | -0.44 | 0.62 |
| Ontustik-Kazakhstan oblysy | 1.88 | 1.22 | -2.03 | -0.89 | -0.16 | -0.07 | 0.31 | -0.98 | 0.25 | 0.11 | 0.73 | -0.09 | 0.83 | -1.18 | 0.51 | -0.44 | 0.00 |
| **Total** | 1.68 | 1.02 | -1.75 | -1.94 | 0.33 | 0.16 | 0.49 | -0.19 | -0.64 | -0.11 | 0.52 | 0.16 | 1.32 | -1.74 | 0.89 | -0.37 | 0.16 |
| **Guinea (2011-2005)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Basse Guinée | -1.00 | -0.36 | 0.95 | 0.11 | 0.84 | 1.07 | -0.43 | -0.80 | -0.31 | -1.78 | 0.59 | 0.03 | 0.35 | 0.07 | 0.45 | -0.05 | 0.29 |
| Moyenne Guinée | 0.33 | 0.40 | -1.42 | -1.17 | 0.67 | 0.74 | -0.14 | -1.03 | -0.46 | -0.17 | 0.46 | 0.10 | 0.54 | 0.20 | -0.03 | 0.37 | 0.64 |
| Haute Guinée | 0.59 | -1.32 | 0.07 | 1.31 | 1.34 | 0.35 | 0.32 | -1.06 | -0.12 | -0.48 | 0.22 | -0.03 | -0.07 | -0.35 | -0.32 | -0.13 | -0.30 |
| Guinée Forestiere | -0.16 | 0.19 | 0.18 | 0.74 | 0.05 | 0.51 | -0.54 | -0.50 | -0.01 | -0.69 | 0.55 | -0.33 | 0.13 | 0.08 | -0.06 | -0.09 | -0.09 |
| Conakry | 0.86 | -1.93 | 0.91 | -0.39 | 0.61 | 0.75 | -0.52 | -0.54 | -0.49 | -0.48 | 0.36 | 0.49 | 0.18 | 0.11 | 0.05 | 0.14 | 0.04 |
| **Total** | -0.04 | -0.62 | 0.18 | 0.22 | 0.82 | 0.80 | -0.22 | -0.78 | -0.27 | -0.79 | 0.43 | 0.00 | 0.21 | -0.02 | -0.01 | 0.01 | 0.10 |
| **Guinea (2011-1999)** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Basse Guinée | -1.71 | -2.17 | 1.50 | -0.17 | 0.63 | 0.00 | -0.85 | -0.33 | 0.24 | -0.73 | 1.21 | 0.21 | 0.75 | 0.25 | 0.52 | 0.37 | 0.33 |
| Moyenne Guinée | 0.75 | -1.64 | -1.34 | -0.33 | -0.06 | -0.07 | -0.21 | -0.62 | -0.23 | 0.70 | 0.63 | 0.45 | 0.32 | 0.51 | 0.03 | 0.62 | 0.53 |
| Haute Guinée | 0.36 | -0.31 | 2.24 | 0.14 | -0.19 | -0.60 | 0.06 | -0.60 | 0.32 | -0.14 | 0.43 | -0.37 | -0.50 | -0.19 | -0.18 | -0.30 | -0.11 |
| Guinée Forestiere | -0.65 | -0.16 | 1.55 | 0.89 | 0.06 | -0.66 | -0.57 | -0.66 | 0.47 | -0.73 | 0.52 | -0.18 | -0.06 | 0.04 | 0.22 | 0.00 | -0.02 |
| Conakry | 0.30 | -2.36 | 1.31 | 0.00 | 1.29 | -0.78 | -0.97 | -0.57 | 0.03 | -0.87 | 0.95 | 0.49 | 0.92 | 0.09 | 0.04 | 0.25 | 0.11 |
| **Total** | -0.18 | -1.41 | 0.98 | 0.24 | 0.44 | -0.27 | -0.44 | -0.53 | 0.17 | -0.35 | 0.67 | 0.08 | 0.21 | 0.09 | 0.09 | 0.16 | 0.13 |

Table 5. *Absolute change (difference) in age group specific percentage points between survey round years in Nepal (DHS Surveys 2011-2006 & 2011-2001), Kazakhstan (MICS Surveys 2011-2006), Guinea (DHS Surveys 2011-2005 & 2011-1999), for each region and total country. Notes: percent difference is given as %-points. DHS = Demographic and Health Survey; MICS = Multiple Indicator Cluster Survey.*