

Interim Report

IR-12-016

Populations for 171 Countries by Age, Sex, and Level of Education around 2010: Harmonized Estimates of the Baseline Data for the Wittgenstein Centre Projections

Ramon Bauer (ramon.bauer@oeaw.ac.at)
Michaela Potančoková (potancok@iiasa.ac.at)
Anne Goujon (goujon@iiasa.ac.at)
Samir K.C. (kc@iiasa.ac.at)

Approved by

Wolfgang Lutz
Leader, World Population Program

December 20, 2012

Interim Reports on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

Contents

1	Introduction	1
2	Data Sources	2
3	Data Harmonisation	5
3.1.	Categories of educational attainment.....	6
3.2.	Allocation procedures	7
4	Data Validation.....	15
4.1	Different data sources	15
4.2	Consistency within the dataset.....	18
4.3	Comparison with other datasets	19
5	Data Adjustments	19
5.1	Missing educational categories	20
5.2	Age groups adjustments.....	20
6	Comparison between the 2007 and the 2012 Datasets	20
7	Conclusion.....	22
8	References	23
	Appendix Tables.....	24

Abstract

Data series on levels of educational attainment of the adult population consistent across time and space cannot be found as such, readily available, not in an aggregated form and not by age and sex. This is a pity because levels of educational attainment of the working age population are the main component of human capital that is used in many models, mostly related to economics, IT and health. Researchers at IIASA have developed a methodology to reconstruct and project levels of educational attainment (see Lutz et al. 2007) based on the information contained in the best source for the most recent year. An improved and increased version will become available in 2013. We are showing in this paper that it does not really make sense to keep the data as close as possible to those directly available datasets since a large majority of those suffer from severe flaws, hampering any trend and regression analysis on levels of educational attainment. We show how picking the right dataset for the starting year can be a real hassle and point towards the necessity to invest in harmonizing and mapping levels of education to facilitate academic research for the benefit of societies.

Acknowledgments

Numerous people have helped and advised us with data collection and this work would have been of lesser values without their assistance: Mohammad Jalal Abbasi, Talgat Akkbakov, Huda Alkitkat, Mohamed Bedrouni, Gervais Beninguisse , Helge Brunborg, Isabella Buber-Ennser, Beverley Busby, Gui Ying Cao, Ania Chaluda, Youssef Courbage, Aleksandra Danilovic, Paola di Giulio, Danilo Dolen, Sangai Dorji, Regina Fuchs, Alessandra Garbero, Francisco Gonzales, Clarissa Guimaraes, Meimanat Hosseini, Selvata Hot, Leiwen Jiang, Jungho Kim, Nato Kopaleishvili, Mohammed Kouidri, Elke Loichinger, Éric Caron Malenfant, Ángel de la Fuente Moreno, Robert McCaa and the IPUMS-international project team, Raya Muttarak, Kelebogile Olifant, Ülle Pettai, Dimiter Philipov, Ivanka Purić, Regina Radinger, Fernando Riosmena, Sanda Roze, Nikola Sander, Andreas Sattr, Sergei Scherbov, Riikka Schemeika, Markus Springer, Marcin Stonawski, Maria Rita Testa, Kuenga Tshering, Ásta M. Urbancic, Didier Willaert, Sam Yoo, Hassan Yousif, Brenda Yepez-Martinez, and Krystof Zeman. Thanks also to all of those “anonymous” working in statistical office, who answered positively our request for data. Funding for this work was made possible by the European Research Council (ERC) Advanced Investigator Grant focusing on “Forecasting Societies’ Adaptive Capacities to Climate Change” (ERC-2008-AdG 230195-FutureSoc).

About the Authors

Ramon Bauer is Research Scientist at the Vienna Institute of Demography (VID) of the Austrian Academy of Sciences, Wittgenstein Centre for Demography and Global Human Capital.

Michaela Potančoková is Research Scientist at the Vienna Institute of Demography (VID) of the Austrian Academy of Sciences and Research Scholar in the World Population Program at the International Institute for Applied Systems Analysis (IIASA), Wittgenstein Centre for Demography and Global Human Capital.

Anne Goujon is Leader of Research Group “Human capital and migration” at the Vienna Institute of Demography (VID) of the Austrian Academy of Sciences and Senior Research Scholar in the World Population Program at the International Institute for Applied Systems Analysis (IIASA), Wittgenstein Centre for Demography and Global Human Capital.

Samir KC is Leader of the “Modeling Human Capital Formation” project in the World Population Program at the International Institute for Applied Systems Analysis (IIASA), Wittgenstein Centre for Demography and Global Human Capital.

Populations for 171 Countries by Age, Sex, and Level of Education around 2010: Harmonized Estimates of the Baseline Data for the Wittgenstein Centre Projections

Ramon Bauer
Michaela Potančoková
Anne Goujon
Samir K.C.

1 Introduction

In 2007 and 2008, International Institute for Applied Systems Analysis (IIASA) and Vienna Institute of Demography (VID) published the first round of global population projections by level of educational attainment from 2000 to 2050 as well as the reconstruction – also called back-projections – back to 1970 (Lutz et al. 2007, K.C. et al. 2008 and K.C. et al. 2010). Both the projections and back-projections – using the cohort-component method adopted for multistate projections – required base-year data on population disaggregated by levels of educational attainment by age and sex. The IIASA/VID 2007 dataset included 120 countries around the year 2000 by four education categories. In 2011, in the framework of the Wittgenstein Centre for Demography and Global Human Capital (WIC)¹, it was decided to implement a new round of projections for three main reasons: (1) update the previous round with more recent data from censuses and surveys; (2) increase the number of education categories from four to six to encompass more differentials in levels of attainment across the world; and (3) increase the number of countries to be able to draw a global vision of levels of educational attainment and their potential future (Lutz et al. 2013). For this purpose, there was a need to update and extend the IIASA/VID 2007 data collection of educational attainment by coverage and level of detail, which is the exercise that is presented in this paper.

Our aim was to collect data on population by age, sex and educational attainment for 195 countries with a population of at least 100,000 listed in the 2010 Revision of the UN World Population Prospects². However, it was a challenge to obtain data for all these countries in sufficient detail. In total we managed to collect and harmonise data on educational attainment by age and sex for 171 countries (88 % of all countries), covering 97.4 % of the world population in 2010. The countries coverage has significantly improved compared to the previous round of projections published in 2007, which covered 120 countries with at least 100,000 inhabitants. In spite of our efforts, it was not possible to include 24 countries (see Table 1), because recent data on educational attainment were either not available at all (e.g. North Korea, Papua New Guinea), or the data at hand were not at the sufficient level of detail or quality (e.g. Angola, Afghanistan). In general, human capital

¹ The Wittgenstein Centre for Demography and Global Human Capital (WIC) was established in 2011 as a collaborative effort between the World Population Program of IIASA, the Vienna Institute of Demography of the Austrian Academy of Sciences (OEAW) and the Research Institute on Human Capital and Development of the WU (Vienna University of Economics and Business).

² We based our list on the countries recognised by the UN in 2010 (http://esa.un.org/wpp/Excel-Data/WPP2010_F01_LOCATIONS.XLS). Therefore, countries such as Kosovo, Taiwan and South Sudan (which gained independence in July 2011) are not included in the WIC 2012 dataset.

stocks are much more difficult to obtain than data on actual education flows such as school attendance, completion or drop-out rates. A lot of information is available for the population at school age or those who were in education or training at the time of the census or survey, but very little on the educational composition (i.e. highest level attained) of the entire population. This is quite amazing, given that evidence on human capital stocks by educational attainment contains valuable information about social change that is highly relevant for human capital research and policy-maker.

At first glance, the undertaking seems to be simple enough: to collect data on highest level of educational attainment for as many countries as possible, which should be available from recent censuses and surveys. As the following sections will show, the task is not as trivial as one might think because such data are very rarely readily available. First, Section 2 discusses the WIC data collection with respect to data availability and the particularities of various data sources. In Section 3, we present the WIC 2012 categories on educational attainment and describe the challenges connected to the allocation of national categories into our six categories. Section 4 addresses the validation of the WIC 2012 dataset and Section 5 summarises the necessary data adjustments that were carried out. Section 6 analyses the differences between the new WIC 2012 dataset and its predecessor, the IIASA/VID 2007 dataset on population by educational attainment. The concluding section summarises the main features and envisaged applications of the new WIC 2012 dataset and discusses the next steps to improve this unique data collection on global human capital.

2 Data Sources

The efforts of collecting and harmonising data on various populations by education undertaken under the WIC 2012 round of global human capital projections are neither the first nor the only efforts in this field. However, most of the collections of data on education are related to flows in the schooling systems (enrolment, repetition, transition, completion) rather than stocks of education that are the translation of the flows in terms of levels of educational attainment of the adult population. As to the collection of data on educational attainment, they are of three major types. The first one is collected by major international institutions such as the United Nations Educational, Scientific and Cultural Organization (UNESCO) or EUROSTAT directly from the statistical offices. The data suffer from the flaws in the reporting from the statistical institutes which international bodies have to take at face value. The second type of collection is based on those collected at the sources such Demographic and Health Surveys (DHS) or MICS. Beside the problems of sampling that will be addressed later in the paper, they tend to be designed for the national context and are not always immediately comparable. The third type is close to the exercise undertaken here and aims at collecting data on levels of educational attainment from many sources. Examples are numerous: Education Policy Data Center (EPDC), Barro and Lee, IIASA/VID. Most of these datasets are also not satisfactory since they also tend to take the value collected elsewhere at face value.

In order to collect the most reliable and up-to-date data on population by age, sex and educational attainment, we examined various data sources with a special emphasis on detailed education and age categories for the population 15 years and older. In terms of a detailed representation of age we were targeting single or 5-year age groups, from age 15 onwards, at best up to age 100 years and older. With regard to detailed educational categories, we aimed to collect data at a level of detail that ensures a clear allocation to the six WIC 2012 categories (see Table 2). For that reason, we gathered data on both the highest level attained and highest grade attended and school year whenever possible. Such a level of detail enabled us to disentangle the latent ambiguity between completed and incomplete levels of educational attainment. Based on these essential principles, the following hierarchy of potential data

sources emerged. In the first place, we were looking for register or census data, which usually comply with all requirements. If no register data or recent census data from the 2000 or 2010 round were available, we tried to collect data from extensive and representative sample surveys. In some cases, it was necessary to draw on (demographic, labour force or household) surveys with more restricted samples.

Register data are in general reliable, accurate and up to date and hence ideal for our needs. However, very few countries, even in the more developed world, have population registers we could rely on. Therefore, our data collection efforts primarily focused on the census data. National censuses are, in general, an accurate source of valid information on education attainment, but are usually conducted only once in a decade and are hence often outdated. We collected census data from various sources, retrieving data from databases and websites of national statistical offices (NSO) and also from online data sources like the Integrated Public Use Microdata Series (IPUMS) provided by the Minnesota Population Center, CELADE (CEPAL's population division) or EUROSTAT. Whenever possible, we used micro data (as provided by IPUMS) or databases that enabled us to extract custom tables from full census data (such as the CELADE database that provides most Latin American and Caribbean censuses). In case we could not find detailed census reports or online databases, we requested the raw data directly from national statistical agencies. This endeavour was quite successful in many countries, but turned out to be rather fruitless in the developing world, especially in sub-Saharan Africa. In general, the level of detail of census data provided by national statistical offices varied tremendously. Sometimes we came across data that did not differentiate by sex, or we just found tables on literacy or school attendance. Quite often, standard tables on educational attainment were aggregated for the entire population above certain age. Lacking the age dimension is naturally a problem since educational attainment differs substantially among cohorts, especially in countries that have recently experienced a significant expansion in education.

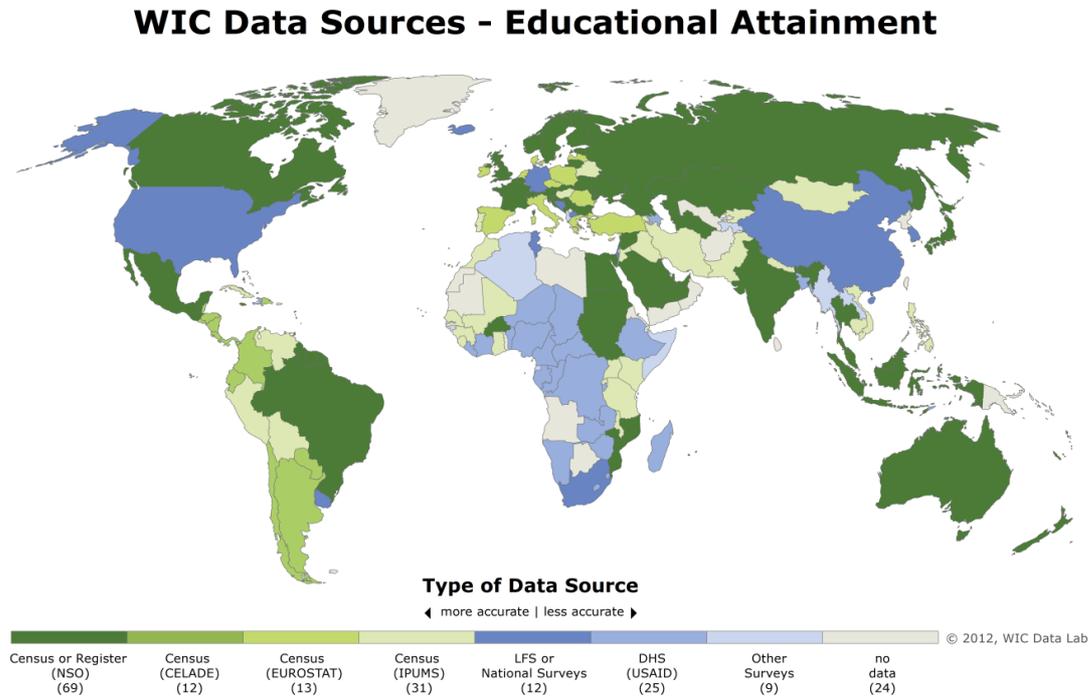
In those cases when census data were of poor quality, not available or outdated, we turned to surveys in order to capture the educational composition of the population. This was mostly the case in developing countries, even so a few European countries did also not conduct a census in recent decades. Preferably, we targeted extensive and representative sample surveys as reliable sources of information on educational attainment, such as national Labour Force Surveys (LFS) for Germany and Bosnia and Herzegovina or other large-scale surveys like the American Community Survey (ACS) for the USA or the National Population Sample Survey for China.

In the absence of extensive national surveys, we had to turn to household surveys on demographic, health and socio-economic issues. This was particularly the case in many African countries, where we used primarily Demographic and Health Surveys (DHS). If no DHS has been carried during the past 15 years or if a particular survey wave was lacking information on educational attainment (of all household members), we were looking at other surveys like the Multiple Indicator Cluster Survey (MICS) for Burundi, Gambia, Equatorial Guinea, Guinea-Bissau and Somalia, the Pan Arab Family Health Survey (PAPFAM) for Algeria, the World Bank's Living Standards Measurement Study (LSMS) for Albania and Tajikistan, Regional Health Surveys (RHS) in Lao and Myanmar, or other household surveys.

It must be noted that some of these surveys do not ensure a fully representative sample across gender, age and territory. Labour force surveys tend to overstate educational attainment as those in labour force tend to be more educated (Barro and Lee 2001). Household surveys on demographic and health issues like the DHS focus on women of reproductive age, and in particular the poorest households or those located in remote areas are more likely to be

omitted from or refuse to participate in surveys³. Consequently, data based on such surveys were treated with due care to avoid biased results with respect to the educational composition of populations in countries concerned.

Figure 1: WIC 2012 data sources on educational attainment (effective December, 2012).



³ Coverage is an issue not only for surveys, but also for some censuses. For example, the Sudanese census of 2008 covered only a small fraction of the population of the provinces that are now South Sudan.

Table 1: Country coverage of the WIC dataset by UN region

UN region	All countries	Countries covered	Countries covered (%)	Population covered (%)	Missing countries
Europe	39	39	100	100	
Asia	50	43	86.0	96.9	Afghanistan, Brunei, North Korea, Oman, Sri Lanka, Uzbekistan, Yemen
Africa	55	46	83.6	95.7	Angola, Botswana, Djibouti, Eritrea, Libya, Mauritania, Mayotte, Togo, Western Sahara
Northern America	2	2	100	100	
Latin America	37	34	91.9	98.9	Barbados, Grenada, Virgin Islands
Oceania	12	7	58.3	75.8	Fiji, Papua New Guinea, Solomon Islands, Micronesia, Guam
World	195	171	87.7	97.4	

Note: Data collection focus on countries with population of at least 100 000 in 2010.

While it has been challenging for some countries to get hold of any data at all, for others we collected several types of data from various sources and decided on which one to use based on the completeness and accuracy regarding our data need and the data hierarchy above mentioned. In sum, we used register data for Austria and three Nordic countries (Finland, Norway and Sweden). For another 125 countries we have collected census data from the 2000 (96 countries) and 2010 census rounds (29 countries) (see Appendix table 1). For most countries 2010 round census data were not yet released at the time when the data collection process has been completed (November 2012). Nevertheless, it was possible to include census data from the 2010 round for a few populous countries such as Brazil, Indonesia and Japan. That means that we could rely on census data in most cases. If census data were not available, reliable or of disputable quality (e.g. Nigerian census 2006), we turned our attention to representative sample surveys (12 countries), DHS (25 countries) or other household surveys (9 countries).

3 Data Harmonisation

According to the United Nations (UN 2007), educational attainment is defined “*as the highest grade completed within the most advanced level attended in the educational system of the country where the education was received. Some countries may also find it useful to present data on educational attainment in terms of highest grade attended (...). For international purposes, a "grade" is a stage of instruction usually covered in the course of a school year*”.

It is apparent that educational categories describing the highest level of attainment are always based on national educational programmes. Due to the variety of nationally distinct educational systems, many different types of educational levels exist around the globe. In order to make education statistics comparable across countries, UNESCO designed the International Standard Classification of Education (ISCED) already in the early 1970's. Later, this classification had been updated and a revised version ISCED 1997 was implemented (UNESCO 2006 [1997]). The latest revision has been released in 2011 and is being implemented since 2012. However, country-specific ISCED 2011 mappings are not yet released. Therefore and for reasons of comparability – by now, most international

organisations (like OECD, World Bank, Eurostat) and research institutions dealing with cross-national educational statistics are still using the 1997 revision, which literally became the global standard for the collection and analysis of international data on education - the categories used in the WIC 2012 dataset on global educational attainment are not only based on ISCED 1997, but also in line with the new ISCED 2011-levels⁴.

3.1. Categories of educational attainment

In order to harmonise the collected data, first we collected detailed information based on national education categories – optimally categories as surveyed at censuses or surveys without being further processed or aggregated – to allocate the various categories to comparable ISCED 1997-levels, before creating the WIC 2012 categories in a second step. Special emphasis has been placed on the differentiation between completed and incomplete levels, using information about the highest school year or grade attended within the level to distinguish between completed and incomplete levels.

ISCED 1997 differentiates between six main levels of education and additional sub-categories describing the orientation of a particular programme (i.e. general or vocational). These levels are closely related to graduations in national educational programmes and complemented by additional criteria such as starting age, duration and entrance requirement (UNESCO 2006 [1997]). In a nutshell, the main categories of ISCED 1997 can be characterised as follows:

- ISCED 0: Pre-primary education – Preceding primary education (i.e. before compulsory education).
- ISCED 1: Primary Education – In accordance with basic/elementary education (in general compulsory).
- ISCED 2: Lower Secondary (or Second Stage of Basic) Education – Completion of basic education (compulsory in some countries).
- ISCED 3: Upper Secondary Education – A more specialised education prepares participants either for entry into specific occupations or gives access to higher and tertiary education (compulsory in only a few countries).
- ISCED 4: Post-Secondary Non-Tertiary Education – Captures programmes that straddle the boundary between upper secondary and post-secondary education; preparing students either direct entry in the labour market or for further tertiary education.
- ISCED 5: First Stage of Tertiary Education – Not leading directly to an advanced research qualification, either research specific (5A) or technical-occupational (5B).
- ISCED 6: Second Stage of Tertiary Education – Leading to an advanced research qualification (e.g. PhD or Doctorate).

We used ISCED 1997 mappings published by the UNESCO Institute of Statistics (UIS)⁵ to allocate country-specific education categories from censuses and surveys into ISCED 1997 levels. These mappings are available for a vast majority of countries, but there are still some exceptions such as Myanmar, Nicaragua or Costa Rica. In such cases we used

⁴ The WIC 2012 categories are also compatible with ISCED 2011. The main change between ISCED 1997 and ISCED 2011 is the expansion of categories covering tertiary education, from two to four levels. Since the WIC 2012 category “post-secondary” comprises all tertiary levels (see Table X.2 below), it remains in line with ISCED 2011.

⁵ <http://www.uis.unesco.org/Education/ISCEDMappings/Pages/default.aspx>, last visited 30.4.2012

country-specific information provided by the UNESCO International Bureau of Education⁶ on existing levels, number of grades within each educational level and theoretical duration to complete each level of education to assess these previously undesignated categories and assign them to existing ISCED 1997 levels.

The six WIC 2012 categories of educational attainment are strongly linked to ISCED 1997 (and also ISCED 2011) although they are not necessarily congruent. Unlike ISCED 1997, the WIC 2012 categories differentiate between “no education”, “incomplete primary” and “completed primary”, while the WIC 2012 category “post-secondary” comprises three different ISCED levels (4, 5 and 6). The main difference between the previously used VID/IIASA 2007 categories and the new WIC 2012 categories is the increased number of levels – see Table 2 for comparison of ISCED 1997 with WIC 2012 and VID/IIASA 2007 categories.

Table 2: Categories of educational attainment

WIC 2012	ISCED 1997	VID/IIASA 2007
No education	No level or ISCED 0 Grade 1 of ISCED 1 not completed	No education
Incomplete primary	Incomplete ISCED 1	Primary
Completed primary	Completed ISCED 1 Incomplete ISCED 2	
Completed lower secondary	Completed ISCED 2 Incomplete ISCED 3	Secondary
Completed upper secondary	Completed ISCED 3 Incomplete ISCED 4 or 5B	
Post-secondary	ISCED 4 & 5B {first diploma, shorter post-secondary courses} ISCED 5A & 6 {longer post-secondary courses, post-graduate level}	Tertiary

3.2. Allocation procedures

Allocating national categories from censuses and surveys to the corresponding ISCED 1997 levels and creating six WIC 2012 education categories was not straightforward. Our objective was to differentiate between completed and incomplete levels and to get as precise estimate of the highest attained level of education as possible. Consequently, we aimed to collect data that preferably specify both level and highest grade or school year attained within a level and provide information on the highest degree or diploma earned for post-secondary education. In practice, we targeted census data using national education categories and avoided, whenever possible, recoded data such as specific IPUMS categories or EUROSTAT data that is aggregated to groups of several ISCED levels.

Of course, not all countries and data sources provide data of sufficient level of detail. Checking the phrasing of surveyed questions on education was crucial for the accurate assessment of the highest level of education. For example, some surveys (MICS, LSMS) implemented very general questions and some country-specific questionnaires just ask a single question on the highest level attained, while others distinguish between those currently attending and those who have ever been to school. Sometimes only the highest level attended

⁶ World Data on Education 7th or 6th edition if information was not available in the more recent edition: <http://www.ibe.unesco.org/en/services/online-materials/world-data-on-education/seventh-edition-2010-11.html>, resp. <http://www.ibe.unesco.org/Countries/WDE/2006/index.html>

was indicated but no grade or school year, although the information has been surveyed. In such cases we downgraded these persons to the next lower level to make sure that we actually capture the highest level attained. To give an example: If a surveyed person indicated primary education with an unknown number of grades attended or completed at that level, this person's highest level attained has been classified as incomplete primary education in the WIC 2012 dataset – for more information on the WIC 2012 allocation rules, see Table 3 below.

Table 3: Allocation rules for education categories

WIC 2012 categories	Allocation rules
No education	Illiterate persons; persons who have never attended school; persons who were attending 1st grade of primary education at time of survey; persons who have completed 0 years/grades at primary level (ISCED 1); persons attending adult literacy courses at time of survey; persons indicating ISCED 0 as highest educational level; khalwa (lowest level of traditional koranic schools)
Incomplete ISCED 1	grades/years of primary education below the grade of graduation from ISCED 1; persons who completed adult literacy courses; persons attending last grade of ISCED 1 at time of survey; persons who have indicated unknown number of grades/years at ISCED 1 level; traditional koranic schools above khalwa level
Completed ISCED 1	completed last grade of ISCED 1 level; completed grades below the last grade of ISCED 2 level; persons attending last grade of ISCED 2 at time of survey; persons who have indicated unknown number of grades at ISCED 2 level
Completed ISCED 2	completed last grade of ISCED 2 level; completed grades below the last grade of ISCED 3 level; persons attending last grade of ISCED 3 at time of survey; persons who have indicated unknown number of grades at ISCED 3 level
Completed ISCED 3	completed last grade of ISCED 3 level; completed number grades or years below the standard duration at ISCED 4 or ISCED 5B level; persons who have indicated unknown number of grades at ISCED 4 or 5 level
Post-secondary ⁷	Persons who have completed number of years or grades corresponding to standard duration of ISCED 4 or ISCED 5B programmes; persons with completed post-secondary university or non-university education; persons holding degrees corresponding to ISCED 4, ISCED 5B, ISCED 5A and ISCED 6 levels

⁷ This highest category is rather large since it encompasses non-tertiary and tertiary. However, some frequently used data sources (e.g. DHS) were lacking the level of detail at post-secondary education categories, often not differentiating between shorter (ISCED 4 and 5B) and longer post-secondary education (ISCED 5A and 6).

As summarised in Table 3, we have developed standardised procedures to deal with certain obstacles during the process of allocating educational categories. Nevertheless, collecting and harmonising educational data of 171 countries was still a complex task and a challenging exercise, especially because of (a) discrepancies between ISCED categories and those in censuses or surveys, which complicated the attribution to a particular ISCED 1997 category; (b) the assignment of country and region-specific programmes and curricula, especially categories of religious education; (c) changing national education systems, especially the lack of comprehensive overviews that address and document educational structures and changes over time; and (d) often hardly comparable categories at post-secondary levels.

Discrepancies between surveyed categories and ISCED

Although ISCED 1997 mappings are available for most countries, it was not always clear-cut how to identify and allocate some surveyed categories, if the name of the respective national education programme varied (strongly) from those used in UNESCO's ISCED mappings. In general, census categories on educational attainment are designed in correspondence to national education systems and are not necessarily in accordance with ISCED. National categories may comprise several ISCED levels, which was frequently the case for programmes at ISCED 3 and ISCED 4 levels. For example, in the Russian census as well as in other post-soviet countries, some census categories are applied to both ISCED 3 and 4 levels. Consequently, these fuzzy categories needed to be attributed to one or another ISCED category. This was only possible with a particular knowledge on the duration or type of the programme concerned.

Data on the highest level of education in developing countries, where only a small fraction of the population attained tertiary education, very rarely differentiate between shorter (ISCED 4 and 5B that are practically oriented or occupationally specific programmes) and longer post-secondary education (ISCED 5A that are largely theoretically-based and 6). This was no problem within our framework, since the WIC 2012 category "post-secondary" subsumes all levels and programmes beyond ISCED 3. More problematic was the lacking level of detail in lower education categories. This is predominately the case in more developed countries, where the share of lower educated people is generally rather small. Official statistics in Finland or Austria, for example, pool all persons with lower educational attainment than ISCED 2 into a single category. Although nowadays ISCED 2 corresponds to compulsory education in most OECD countries, in many countries the proportions of lower educational attainment are still significant for older cohorts, but do not become clearly evident when hidden in such broad aggregations. Besides that, less detailed information about lower education categories makes it difficult to capture the educational attainment of (often less educated) immigrants from developing countries – for details on the procedures applied in such cases, see Section 5.1.

The other extreme is census data that contains dozens of education categories. In order to adequately allocate them to the six WIC 2012 categories, we relied on codebooks from national statistical institutes as well as on the knowledge of local experts. Such a plethora of education categories can be found, to name a few, in census data from Vietnam (79 categories in 2009), Bolivia (62 categories in 2001) or Iran (60 categories in 2006). For a further illustration, Table 4 shows the original coding in the Iranian census of 2006, as well as the corresponding WIC 2012 categories.

Table 4: Allocation of educational categories in Iranian census 2006 to ISCED 97 levels

	Census categories	ISCED 97 level (completed)		Census categories	ISCED 97 level (completed)
1	None	None	31	Vocational, 6-years system, degree	ISCED 3
2	Adult literacy program	incomplete ISCED 1	32	Pre-vocational or pre-technical, short	ISCED 3
3	Elementary, grade 1	incomplete ISCED 1	33	Vocational or technical, theoretical track, short	ISCED 3
4	Elementary, grade 2	incomplete ISCED 1	34	Vocational or technical, practical track, short	ISCED 3
5	Elementary, grade 3	incomplete ISCED 1	35	Pre-vocational or pre-technical, long track	ISCED 3
6	Elementary, grade 4	incomplete ISCED 1	36	Vocational or technical, long	ISCED 3
7	Elementary, grade 5	ISCED 1/incomplete ISCED 1*	37	Other vocational/technical or teacher training	ISCED 3
8	Elementary, grade 6 (old system)	ISCED 1	38	College, no degree	ISCED 3
9	Unknown grade	incomplete ISCED 1	39	College, degree	ISCED 5
10	Middle school, grade 1	ISCED 1	40	Graduate program, no degree	ISCED 5
11	Middle school, grade 2	ISCED 1	41	Graduate program, degree	ISCED 5
12	Middle school, grade 3	ISCED 2	42	Medicine, no degree	ISCED 5
13	Middle school, unknown grade	ISCED 1	43	Medicine, degree	ISCED 5
14	Middle school (old system), grade 7 or 8	ISCED 1	44	Community college, no degree	ISCED 3
15	Middle school (old system), grade 9	ISCED 2	45	Community college, degree	ISCED 5
16	High school, grade 1	ISCED 2	46	Doctoral program, no degree	ISCED 5
17	High school, grade 2	ISCED 2	47	Doctoral program, degree	ISCED 6
18	High school, grade 3	ISCED 3	48	Postdoctoral program, no degree	ISCED 6
19	High school, diploma	ISCED 3	49	Postdoctoral program, degree	ISCED 6
20	High school, unknown grade	ISCED 2	50	College, unknown	ISCED 5
21	Pre-university, no degree	ISCED 4	51	Shiism theology, 1st level	ISCED 5
22	Pre-university, degree	ISCED 4	52	Shiism theology, 2nd level	ISCED 5
23	Preliminary theology	ISCED 3	53	Shiism theology, 3rd level	ISCED 5

	Census categories	ISCED 97 level (completed)		Census categories	ISCED 97 level (completed)
24	Vocational, credited system, no degree	ISCED 2	54	Seminary school theology, 1st level	unknown
25	Vocational, credited system, degree	ISCED 3	55	Seminary school theology, 2nd level	unknown
26	Training and mastership, credited system, no degree	ISCED 2	56	Theology, Sunni	unknown
27	Training and mastership, credited system, degree	ISCED 3	57	Theology, other sects	unknown
28	Vocational, 4-years system, no degree	ISCED 3	58	Theology, other religions	unknown
29	Vocational, 4-years system, degree	ISCED 3	59	Theology, unknown religion	unknown
30	Vocational, 6-years system, no degree	ISCED 3	60	Unknown	unknown

Notes: *incomplete ISCED 1 has been applied for persons born before 1970 that were enrolled in older education systems before recent reforms were adapted. Categories were recoded using a codebook from the Iranian national statistical office.

Religious education

Categories of religious education in the Iranian census (as shown in Tab X.3 above) reveal another allocation problem, especially when religious schooling exists next to a public school system. Religious education may provide education at all kinds of ISCED levels, from pre-primary to post-secondary education. It was a particular challenge to allocate Koranic or Buddhist schools and educational programmes to the WIC 2012 categories. At Koranic schools, for example, which are common in many Arabic, Maghreb and sub-Saharan countries as well as in Southern Asia, contents and standards of educational programmes vary strongly. On the one side, traditional Koranic schools mainly focus at memorising the Koran. Sometimes students are also taught to read, write and basic numeracy, but often at different stages of schooling (Easton and Peach 1997). On the other side, modern Koranic schools called *médersa* can have curricula similar to public schools. When recoding such specific cases, we used studies describing those curricula or evaluating the quality of religious schools (for example Andre and Demonsant 2009, Easton and Peach 1997, Bledsoe and Robey 1986, El Sammani et al. 1985) and followed the advice given by experts with country-specific knowledge.

Changes in national educating systems

Transformations in educating systems over time were yet another challenge for determining educational attainment by age. People of different ages often went through different educational programmes or systems with respect to the number of grades required to reach a certain ISCED level. Besides prolongations in the duration of compulsory education, alterations in the duration of schooling might occur just at one particular level or at more levels at once. Some countries did change their educating systems fairly frequently. Since 1970, for example, Cambodia had four different systems and Mozambique and Ukraine reformed their educational systems three times. Moreover, such changes in national educating systems are often poorly documented and are thus difficult to identify.

Aiming to take changing education systems into account when allocating original data into the six WIC 2012 categories, we benefited from a compendium of documented changes

compiled by UNESCO⁸. However, UNESCO does not provide any information before the 1970s and detailed information - duration of compulsory education, theoretical duration of ISCED 1997 levels as well as starting age of education at each level - is given only from 1998 onwards. For all persons enrolled prior to 1998, which is almost the entire population of 15 years and older covered by the WIC 2012 dataset, there is only information on the aggregate duration of secondary education, without any differentiation between upper (ISCED 2) and lower (ISCED 3) secondary education. This is a particular problem, if the cumulative duration of lower and upper secondary education remains the same, but the duration of each level has changed - say, from a 3-4 year system to 4-3 years. Such a situation is not unusual, because compulsory education is often extended by adding one or more grades or schooling years to lower secondary education (ISCED 2), often by clipping it off from upper secondary (ISCED 3). In other cases, there was evidence that the cumulative durations of secondary education changed, but no further information about the levels or grades affected was available. Due to these circumstances, we acknowledge that at least some changes in educational systems, which potentially would have affected the cohort-wise allocation of particular grades and degrees to ISCED 1997 levels, may have remained concealed to us.

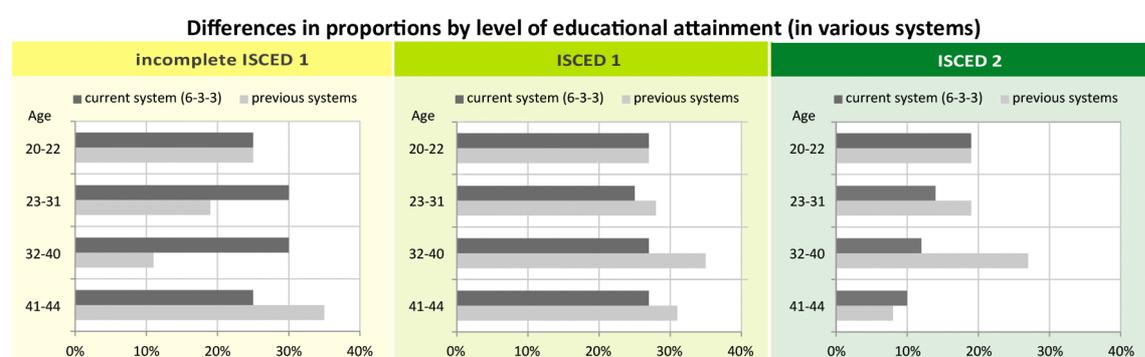
Finding information about changes in national educating systems over time is a big step forward with respect to verifying the actual educational attainment of older cohorts. However, the highest level of education achieved is often surveyed according to the educating system that is effective at the time of the census or survey. Ideally, categories on educational attainment consider older education system by (simply) including separate categories, as it was done in the Iranian census of 2006 (see Table 4). More often, only the current system was taken into account. If sufficient information about educational changes was at hand, we adjusted the educational attainment of corresponding cohorts to the system in which they were actually enrolled (see Appendix Table 1). Sensitivity analyses showed that it makes a lot of difference when the original data was adjusted, as demonstrated by Figure 2 using the example of Cambodia.

⁸ UNESCO institute for Statistics (UIS) collects evidence on past educational systems since 1970. This information can be found online at: http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=143&IF_Language=eng [last visited April 2012]

Figure 2: Differences in educational attainment resulting from changes in the educational system of Cambodia (according to census 2008)

Differences in mandatory grades by level of educational attainment (in various systems)

Age	incomplete ISCED 1					ISCED 1			ISCED 2			ISCED 3	Education system	
	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12		
20 21 22													6-3-3 years (since 1997)	
23 24 25 26 27												5-3-3 years (1986-1996)		
28 29 30 31														
32 33 34 35 36 37 38 39 40														
41 42 43 44														4-3-3 years (1979-1985)
													6-4-3 years (before 1979)	



Note: The age group 20-22 years was not affected by any changes in the education system and serves as a reference group.

At the time of the 2008 census, persons aged 20 to 44 years were enrolled in four different educating systems, including changes both at the primary and secondary level. Before 1979, Cambodia adopted a French-based education system that required thirteen years of education to complete ISCED 3 level (6-4-3). After 1979, the Ministry of Education launched a ten-year education system (4-3-3) and then expanded it to an eleven-year education system during the period 1986 to 1995 (5-3-3). Since the school year 1996/97, the education system was further expanded to twelve years (6-3-3)⁹. The example of Cambodia shows that educational attainment is much lower for persons that were actually enrolled in an older system with shorter primary education¹⁰, if the current education system is applied.

Contrary to some censuses, which at least occasionally take into account older educating systems, surveys are usually solely based on the system that was effective at the time of the data collection. Taking DHS as an example, changes in education systems are only occasionally mentioned in the otherwise extensive documentations. In general, DHS data are recorded accordingly to the educational system at the time of survey. Nevertheless, using information provided by UNESCO, we adjusted the levels of attainment for corresponding cohorts (see Figure 2 showing the example of the Cambodian census in 2008, but also for Jordan 2004 and Argentina 2001). In case we had no information about alternations of the

⁹ See: <http://www.ibe.unesco.org/Countries/WDE/2006/index.html> [last visited April 2012].

¹⁰ We assumed all persons entered schooling at the minimum entrance age and that it took them exactly the theoretical years of schooling to complete a particular level. We are aware that small proportions of cohorts may have been enrolled at different ages, or that it took a longer time to complete a level, or that some even started studying only in adulthood.

educational system, the current system was used to allocate educational categories according to the WIC 2012 scheme, or the respective data source was omitted altogether¹¹. It should be noted that these adjustments explain deviating results in the WIC 2012 dataset when compared to various DHS country reports. In fact, the lack of knowledge of changes in education systems considerably hampers harmonisation exercises such as the efforts undertaken in the WIC 2012 project. Certainly, more detailed and comprehensive information of changes in national education programmes and systems is needed to improve the accuracy of estimating global human capital stocks, especially within older cohorts. However, we believe our more detailed reconstructions – see also Section 5 for data adjustments - improved the quality of the assessment of global educational attainment by age and sex.

Categories at post-secondary level

The highest educational category in the WIC 2012 dataset is post-secondary education, which is defined by the completion of an educational programme at the level of ISCED 4 or higher. However, there are substantial differences how post-secondary education has been surveyed in different censuses and surveys. As already mentioned, the ideal original data would distinguish between incomplete and completed level of education by indicating the highest degree attained to allocate the respective programme using UNESCO's ISCED 1997 mappings. Quite often, especially in Latin American censuses as well as DHS, some surveyed persons indicated post-secondary education but no degree, or only the information on the years studied within a programme or within a level were given but no information on the degree. For example, a person that indicated university education, year 1 completed as the highest level of education attained cannot have a bachelor degree, if the minimum duration of the shortest bachelor program in the country is three years. Consequently, we attributed the educational attainment of persons who have studied programmes at ISCED 4 or ISCED 5B levels (i.e. shorter tertiary – see also Table 5) but did not indicate any completed degree, as ISCED 3 level. The same “downgrading” was applied to persons who indicated less years completed within post-secondary education than actually necessary to obtain a degree at ISCED 5B or ISCED 4 level, as well as to persons that were attending ISCED 4 or ISCED 5B educational programmes at the time of survey (see allocation rules in Table 3). As a result, the proportions in the WIC 2012 dataset at the post-secondary and by implication also at the upper secondary level might deviate from other datasets or publications on educational attainment.

¹¹ In Central Asian countries, the coding of education was particularly problematic, especially because of changes of educating systems. In case DHS data did not correspond to information about the particular educational system, we used a different data source for these countries.

Table 5: Allocation of post-secondary education categories from Mexican census 2010 to ISCED 1997

Census category (level and years completed)	ISCED 1997
University undergraduate:	
Year 1	ISCED 3
Year 2	ISCED3
Year 3	ISCED 5
Year 4	ISCED 5
Year 5	ISCED 5
Year 6+	ISCED 5
Not specified	ISCED 3
Master:	
Year 1	ISCED 5
Year 2+	ISCED 5

4 Data Validation

A special emphasis has been placed on the validation of the new WIC 2012 dataset on global educational attainment. Already throughout the process of data gathering whenever two or more sources were available they were validated against each other in order to reveal the more reliable source. In case that only one source was at hand that met the WIC 2012 criteria (see Section 3), it was our concern to validate the data at least at a higher level of aggregation, for instance using population 15 or 25 years and older, which is often available from other data sources, for example basic tables published by the national statistical agencies or UNESCO database. Besides performing case-by-case comparisons using alternative sources, it appeared reasonable to assess the consistency within the WIC 2012 dataset first before validating it against other existing ones.

4.1 Different data sources

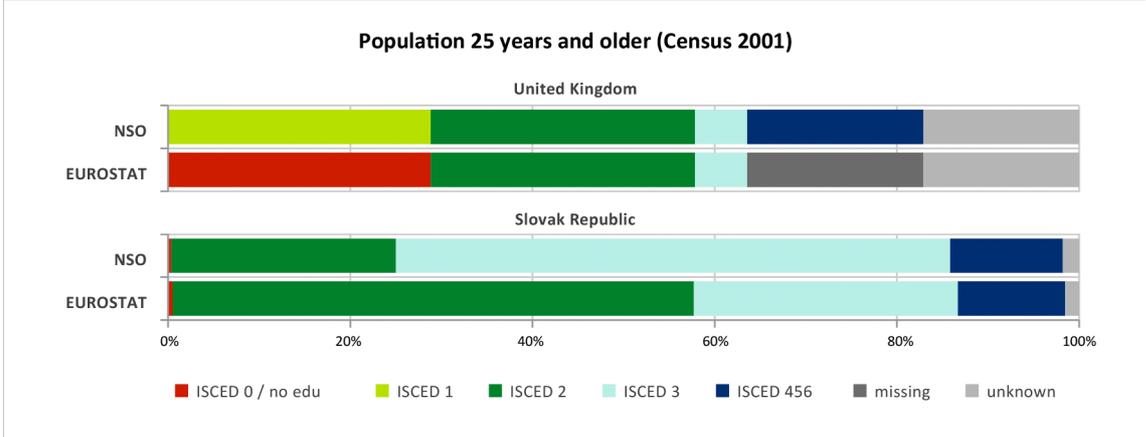
Right from the beginning, we aimed to collect data on educational attainment from various sources in order to choose the most reliable source. In doing so, it became apparent that data sources are both problem and solution when selecting the right source. As a matter of fact, different data sources may lead to different results. In terms of educational data, different sources may result in different educational compositions. Picking the right source is certainly the correct solution to this problem. This was not always an unambiguous task, as the following comparisons between different data sources demonstrate.

National census vs. Eurostat

Generally speaking, register or census data are the best source when collecting information on the highest level of educational attainment. Following this, one would not expect significant differences when comparing census data from different sources. However, if the original data were already aggregated and mapped to conform to ISCED levels, this was not necessarily helpful. Eurostat, for example, provides data from the 2000 census round for 31 European countries on population by sex, age and educational attainment in accordance with ISCED 1997. Although Eurostat validates the data received from the NSOs before

sending it to their database, there are some deviations in Eurostat’s aggregation to ISCED when compared with WIC 2012 recodes of detailed categories taken from national censuses. Figure 3 illustrates such deviations using the example of the United Kingdom and the Slovak Republic.

Figure 3: Different allocation of educational categories in Eurostat and NSO data, UK and Slovak Republic (Census 2001)



In the case of the United Kingdom, Eurostat pooled ISCED 1 (primary education) and ISCED 0 (pre-primary education) into a single category labelled ISCED 0. This is in fact documented in the Eurostat metadata section, but not explained at any point. Furthermore, figures for post-secondary education (i.e. ISCED 4, ISCED 5 and ISCED 6) were not indicated at all in the Eurostat database. According to Eurostat’s documentation, post-secondary education should be included in the “tertiary” section, but such a Eurostat category simply did not exist. It turned out that the total population did not correspond to the sum of all ISCED categories without a remaining residual amount. A comparison of the Eurostat data and the original data from the UK Office for National Statistics (ONS) revealed that this residual amount – indicated as “missing” in Figure 3/top – did exactly match the post-secondary categories of ISCED 4, ISCED 5 and ISCED 6.

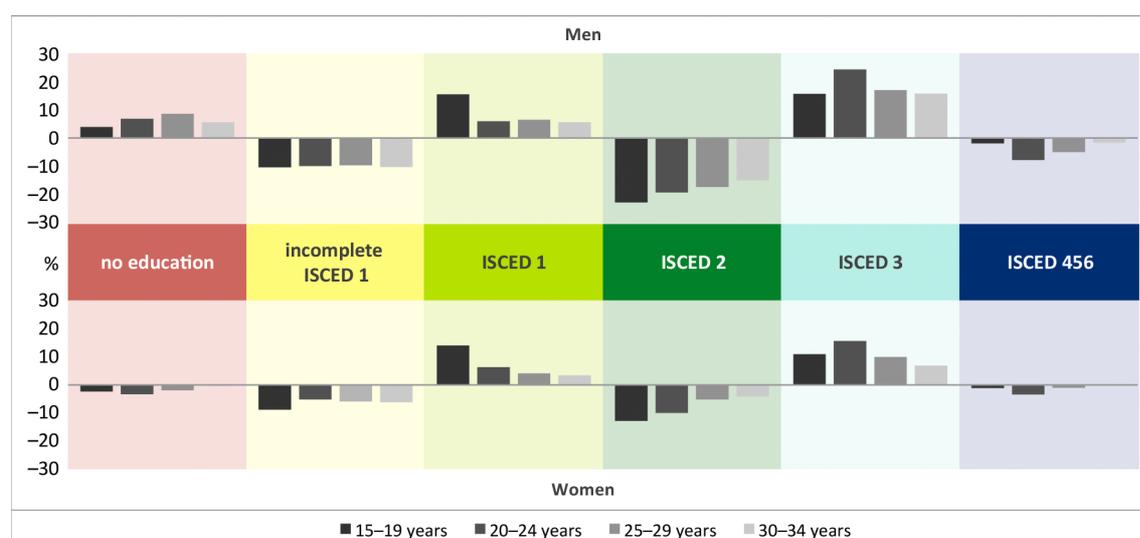
Turning to the example of the Slovak Republic, there are minor differences of less than one per cent in all categories between Eurostat data and those provided by the Statistical Office of the Slovak Republic. More serious, however, is the fact that Eurostat seemed to have overrated ISCED 2 at the cost of ISCED 3. A comparison with detailed educational categories used in the 2001 census of Slovakia revealed that one particular category (that describes a vocational programme without entrance qualification to higher education) was coded to ISCED 2 instead of ISCED 3. As shown in Figure 3/below, this results in a clearly different distribution between lower secondary (ISCED 2) and upper secondary (ISCED 3), which becomes even more significant in older age groups (not depicted in Figure 3). Nevertheless, we used some of the Eurostat data after validating it with NSO data, especially because Eurostat provides data on educational attainment by single-year age groups. This was of great value for additional and more refined analyses in the context of the WIC global human capital projections.

National census vs. DHS

In case no data from recent censuses or other representative surveys were at hand, in particular in African as well as Central and South East Asian countries, information on the highest level of educational attainment was constructed from household or demographic surveys. Above all, the availability of DHS – a household survey that is conducted in over 90 countries and provides data for a wide range of indicators in the areas of population and health, including education – is of great convenience. Nevertheless, sample-based surveys like

DHS yield different results compared to censuses, which aim to cover entire populations. Figure 4 illustrates such differences in the educational compositions of Nepal when comparing census data and DHS data, both from 2001. Although both sources reveal a similar trend, the distributions differ significantly. While the share of no education still remains quite similar in both datasets, the proportions of those with incomplete primary, completed primary (ISCED 1) and completed lower secondary education (ISCED 2) vary substantially – often by a relative difference of more than one hundred per cent. In general, and as demonstrated by the example of Nepal, surveys often tend to underestimate lower education categories and overestimate higher categories, which results in a bias towards a higher level of education.

Figure 4: Absolute difference in percentage points between the educational composition according to the 2001 census and the 2001 DHS in Nepal for selected age groups (census minus DHS)

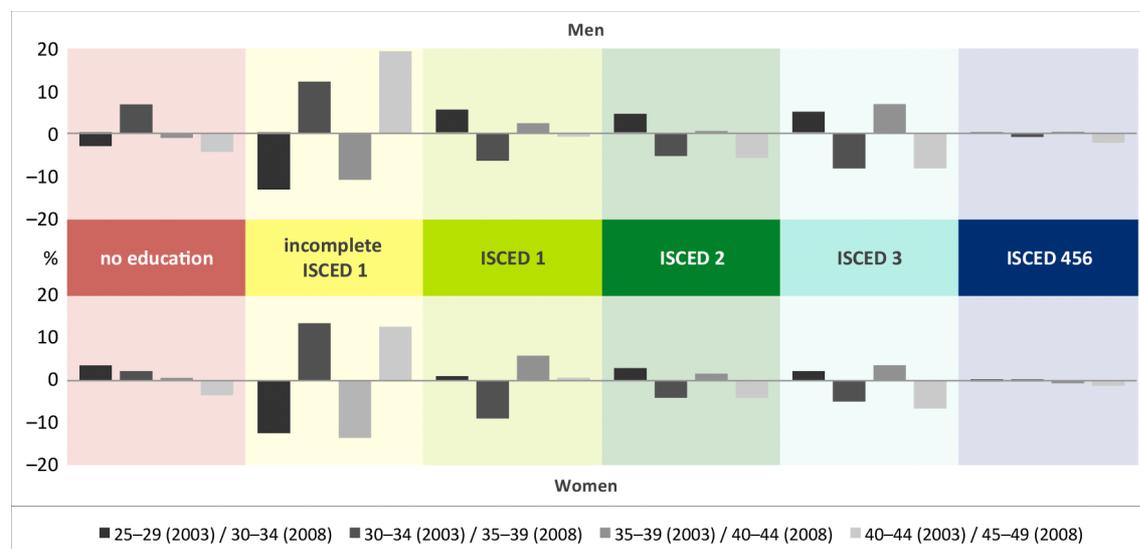


DHS vs. DHS

Results do not only differ by source, even the same type of survey is not necessarily consistent over time, as illustrated by the differences in the educational composition between the 2003 and 2008 DHS rounds in Madagascar. Figure 5 shows the same five-year age group – aged 25 to 44 years in 2003 and 30 to 49 years in 2008 – and the differences in the differences in the educational composition.

Such varieties may be ascribed to differences in the selection of respondents and/or changes in the survey questionnaire. Furthermore, household surveys like DHS often tend to combine completed and incomplete levels of education into a single category when publishing data in their survey reports. For that reason, DHS summary files on population by education were not used for the WIC 2012 database, but rather data that were specifically recoded using DHS microdata and mapped in accordance with the WIC 2012 quality criteria and allocation rules described in Section 3.1. If no DHS dataset was available, restricted or included only incomplete information on educational attainment (e.g. only the educational level of women aged 15 to 49 years), other and – with respect to sample size and strata, territorial coverage or other quality issues – more limited household surveys were used for the WIC 2012 mapping exercise, such as MICS (for several countries in sub-Saharan Africa), LSMS (Albania and Tajikistan), PAPFAM (Algeria) or RHS (Lao and Myanmar).

Figure 5: Absolute difference in percentage points between the educational composition according to the 2008 DHS and the 2003 DHS in Madagascar for selected cohorts (2008 minus 2003)



4.2 Consistency within the dataset

In order to get a first comprehensive overview on the recoded data on global educational attainment, we examined the new WIC 2012 dataset by the simple but effective way of maps. This enabled us to capture the educational level of many countries at a glance – apart from those that are too small to view on the global scale, i.e. a world map. Nevertheless, aiming to detect literally eye-catching discrepancies, we compared the educational compositions of the population 25 years and older of different countries against each other in two different ways: (a) by looking at the share of the lower educated population (with ISCED 1 and lower), which is especially relevant when comparing lower developed countries; and (b) by looking at the share of the higher (post-secondary) educated population, which in turn is relevant when comparing higher developed. In each case the exercise was performed separately for men and women. This intuitive approach indeed revealed some inconsistencies within the dataset, mainly due to the nature of the available ISCED mappings, which we applied when allocating specific national education categories into ISCED 1997 (see Section 3.1). However, one main objective of the WIC 2012 allocation rules is the strict commitment to ISCED 1997. Hence, discrepancies inherent to the general scheme of ISCED 1997 were not adjusted but documented for further considerations – i.e. for comparisons with alternative or more recent sources or additional literature reviews of individual cases in future updates.

A different approach to identify inconsistencies within the WIC 2012 dataset is to examine not only the shares of aggregated educational categories across countries, but also the variation across the six WIC 2012 categories between different countries. Applying principal component analysis (PCA), the countries included in the dataset were classified according to two dimensions (components): (a) in low or high educated countries with (b) more or less variations between the six educational categories. This classification exercise allowed us to identify outlier with respect to suspiciously high concentrations in one or another educational category. It turned out that such concentrations were quite often country-specific particularities in national education systems (e.g. related to different length in compulsory education), or inherent to UNESCO's ISCED mappings of national educational categories. In

some other cases, such outlier indicated educational levels that were either already merged or were not allocated properly.

4.3 Comparison with other datasets

After validating the consistency within the WIC 2012 dataset, the new dataset was also compared against other existing datasets of educational attainment based on ISCED 1997. Since comparisons to datasets that result from reconstruction exercises (see also Section 2) are not necessarily appropriate, we rather concentrated our efforts on other comprehensive data collections. Above all, the UNESCO dataset on educational attainment as published in the Global Educational Digest 2011 (GED 2011) (UNESCO-UIS 2011) was of particular importance. The GED 2011 dataset provides a collection of education statistics for the population 25 years and older by sex that is based on ISCED – i.e. UNESCO's very own classification of international data on education. The UNESCO Institute of Statistics (UIS) compiles educational statistics in aggregate form from official administrative sources at the national level through various surveys¹² and publishes the data – contrary to the WIC 2012 dataset – without any further adjustments. For this reason, and because the UNESCO data is not necessarily based on the same sources or periods as the WIC 2012 data, it turned out that the two datasets are actually hard to compare. In fact, only a few countries were in accordance when checking both datasets against each other. The UNESCO data collection generally tends to show higher educational levels when compared to WIC 2012 dataset, which can be explained (again) by the WIC 2012 approach that aims to distinguish between completed and incomplete levels of attainment in order to downgrade the latter to a lower category.

Other comparisons, for example with data on educational attainment from the UN Statistical Division (UNSD)¹³, did yield similar results. UNSD data are based on national census results and are partially identical with the UNESCO dataset, and hence not as consistent as WIC recoded data in terms of precise distinction between completed and incomplete levels of education. When comparing UNSD and WIC 2012 data for Peru, both based on census data that was recoded from national educational categories to ISCED 1997, it became obvious that UNSD pooled incomplete and completed ISCED levels into one category, while the WIC recode considers only completed levels of education. Hence, it is reasonable that such differences in data allocation and aggregation between UNSD and WIC 2012 result in different outcomes. Due to the particular WIC 2012 approach – i.e. distinguishing between completed and incomplete levels of education, controlling for sensitivity of country-specific educational categories, as well as taking into account changes in national education systems – it is not surprising but rather plausible that it differ from other existing datasets on educational attainment. However, the most crucial question still remains to be answered: what are the differences when comparing the new WIC 2012 dataset with its predecessor, the VID/IIASA 2007 dataset? But before addressing this issue in Section 6, the coming Section 5 outlines the adjustments that were actually applied to the WIC 2012 dataset.

5 Data Adjustments

The WIC 2012 dataset is based on various sources that differ by accuracy and level of detail. For that reason it was inevitable to perform at least a few adjustments to obtain detailed and comparable education and age categories. The adjustments address issues of missing educational categories, age group adjustments and the harmonisation towards a common base

¹² See: <http://www.uis.unesco.org/UISQuestionnaires/Pages/Education.aspx> (last retrieved: 26.05.2012)

¹³ See: <http://data.un.org/Data.aspx?d=POP&f=tableCode%3a30> (last retrieved 02.03.2012)

year. For a full list of adjustments conducted within the WIC 2012 dataset, see Appendix Table 1.

5.1 Missing educational categories

Before dealing with aggregated age groups and base-year adjustments, it proved necessary to resolve the issue of so-called “missing categories”. Actually, there are no educational categories missing in the original data, but occasionally some national categories were either not explicitly indicated or could not be clearly allocated to one ISCED 1997 level or another. In general, official statistics in more developed countries tend to merge several lower educational categories into a single residual group for “primary or compulsory education and lower” (i.e. ISCED 1 or ISCED 2 and lower), while statistics of developing countries – as well as the WIC 2012 dataset – rather aggregate all post-secondary levels into one “higher education” category (i.e. ISCED 4 and higher). Some national categories, however, do not differentiate between lower and upper secondary (ISCED 2 and ISCED 3) or – even more inconvenient - completed or incomplete levels of education.

The best solution to solve such ambiguities is to allocate the categories according to the highest grade attained. If no grades were available in the original data, either additional information from other sources or analogies of populations with similar educational compositions and systems were used to split (already) aggregated categories or to distinguish between fuzzy original categories. It occurred that in some countries one or another educational category needed to be estimated, commonly no education or incomplete ISCED 1. This was normally no problem for developing countries, simply because there are sufficient comparable cases of other countries with similar educational compositions to draw suitable analogies. In more developed countries there is less evidence about the distribution in lower educational categories. Official statistics in many OECD countries tend to pool the lowest educational levels to one category “ISCED 1 or ISCED 2 and lower” – in the extreme case of Finland no category below ISCED 3 was indicated in the original data. Nevertheless, we also estimated “missing categories” for more developed countries whenever suitable analogies were at hand. Due to the WIC 2012 data quality criteria (as described in Section 3), we refrained from any “guesstimation” beyond solid evidence. As a consequence, the WIC 2012 dataset does include a few countries with less than the intended six categories of educational attainment (see Appendix Tables 1 and 2). That concerns only more developed countries with pooled lower categories (e.g. ISCED 1 and lower).

5.2 Age groups adjustments

In some cases interpolations were expedient to estimate 5-year age groups if the original data was only available in another form such as broader or uneven age groups (e.g. 10-year age groups or uneven age groups which are relevant to represent enrolment rates such as 16-18 years or 19-24 years).

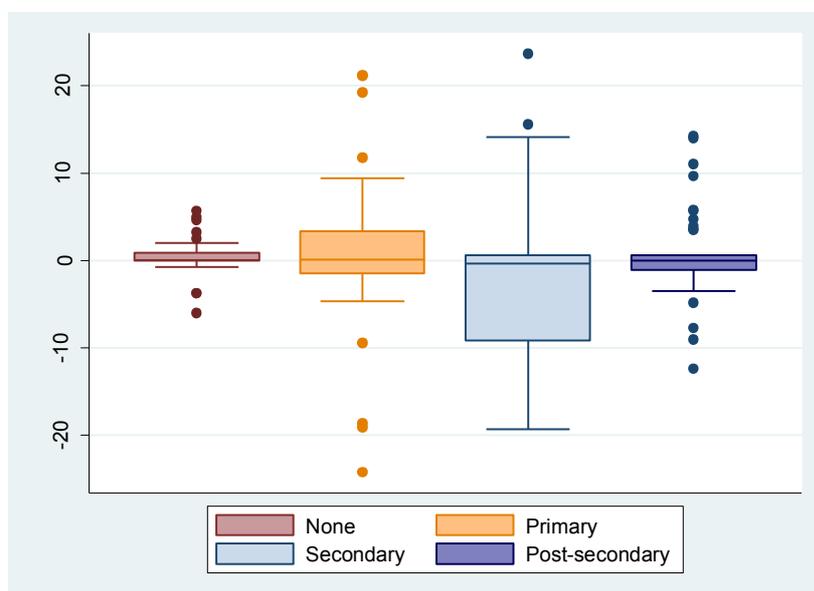
More frequently, trend extrapolations were applied to extend the oldest age group to 100 years and older (100+), because the highest age group is usually an open one (e.g. 65+ or 80+), or the original data did not cover older age groups at all due to restrictions to a certain maximum age (e.g. age 74, as it was the case for some LFS data). In doing so, we follow the procedure described in Section 4.4 of Lutz et al. (2007).

6 Comparison between the 2007 and the 2012 Datasets

The database that was produced in 2007 (see Lutz *et al* 2007 and Riosmena *et al* 2008), which included estimates of the age-, sex- and education specific population distributions of 120

countries, constituted the first effort to create an harmonized database for the projection and reconstruction of future and past levels of educational attainment. With only four education categories (see Table 2), the sampling methodology was very close to that of the present exercise, except that a large part of the data collected relied upon UNESCO data (census data for 35 countries), together with other main sources like DHS (33 countries), and NSO data (28 countries)¹⁴. The stock-figure 6 below shows a comparison between the absolute difference in percentage points between the two datasets for the four categories that were present in the 2007 exercise – conveniently the six categories of the present exercise aggregate themselves very well to the four categories of the previous one – and for the countries with the same base year and type of data. As we have showed in Section 4, comparisons for the countries with different baseline data would necessarily lead to some differences in educational distribution. In total, the comparison is possible for 42 countries. In spite of all problems in harmonisation, for 15 out of 42 countries the correspondence between the two datasets ranges within 3% in each category.

Figure 6: Box plot of absolute differences between 2007 and 2012 datasets by 4 categories of educational attainment



Note: Correspondence between the categories in the two datasets: None = None; Primary = incomplete + completed Primary; Secondary = completed Lower and Upper Secondary, Post-secondary = Tertiary.

It is noticeable that the two databases fare rather well for the two extreme education categories – particularly for the no education one – for which centrality measure are close to 0. Definitions of the lowest and highest education categories are pretty straightforward and, therefore, harmonisation was less problematic for these categories compared to primary and secondary levels. For others, the differences are sometimes quite large: 10 to 20 per cent. This is particularly true for the secondary education due to more precise allocation of the incomplete lower secondary into the primary education as well as allocation of those with incomplete postsecondary education into secondary. Although many efforts in 2007 were put into solving the inconsistency issues between educational attainment categories used in some surveys such as the DHS and the four ISCED based categories, the set of adjustment factors based on the regression of the 10 countries for which recent UNESCO and DHS were available¹⁵, is less reliable than the approach developed under this exercise which through the combined analysis of grades and levels could translate the DHS categories into our categories

¹⁴ Other data sources were from Eurostat (16 countries) and LFS (eight countries)

¹⁵ Armenia, Brazil, Côte d'Ivoire, Guatemala, Jordan, Namibia, Peru, South Africa, Tanzania and Turkey

more precisely. The differences are also due to the choice of the UNESCO source for a large number of countries where the allocation of data to the UNESCO categories, usually performed by the statistical offices, is not very correct, especially in terms of level of completion.

7 Conclusion

Education is a key indicator for appraising the level of socio-economic development of the population in a country. In turn, its measurement can be used for modelling interactions between education and other parameters strongly correlated with education such as, for example, the fertility of women, the capacity of populations to cope with climate change related disasters, economic growth, etc. However, the measurement of educational attainment has always been a problem and despite many attempts to standardize levels of educational attainment, it has not been possible to fully remove all discrepancies across countries in the world, not to mention, across time and age. It would be arrogant to pretend that we have circumvented all obstacles and have created the perfect database on highest level of educational attainment for 171 countries but the efforts which were undertaken in the course of this exercise certainly address the main issues and adopted clear and systematic measures to overcome the failures. Moreover, the strength of the exercise lies also in the exhaustive documentation (see also the appendix tables) of our approach that will facilitate replication and further enhancement. Hence, we are one step closer to the harmonisation of levels of educational attainment of the global population. What remains to be done by national or international organisations, is to enhance the data collection and classification efforts.

The main objective for creating a solid and harmonised dataset on levels of educational attainment is to estimate education distribution of the base year population by age and sex for the new round of multistate population projections by levels of education to be released in 2013. The base year population is one of the main ingredients of the projections besides the three components of population change that are births, deaths and migration – together with changes in educational attainment in the case of multi-educational state population projections. At the time of finishing this report, our dataset contains 171 countries for which we have estimated a distribution into six levels of educational attainment by age and sex. It is worth noting that data is not a perfect reflection of the world, but rather a representation of the world gathered and edited for specific purposes. In case of educational data, the purpose is research on human capital. The WIC 2012 dataset represents the state of the world education according to ISCED 1997. Just like any other classification, also ISCED is a generalisation that cannot reflect all the various details and particularities of the educational systems of every country in the world, as well as the quality of education. However, ISCED is the commonly accepted classification of education, which makes it comparable to a great extent. Certainly, the WIC 2012 dataset is one of the most comprehensive collections of information on global human capital in terms of coverage (97.4% of the world population), sample size (largely based on census data), level of detail (6 categories) as well as accuracy with respect to data harmonisation (systematic approach). We plan to update the dataset regularly and the database as well as the projections will be made available soon on the web at the following address:

www.iiasa.ac.at/web/home/research/researchPrograms/WorldPopulation/POP.en.html

8 References

- Andre, P. and J.-L. Demonsant. 2009. Koranic schools in Senegal: A real barrier to formal education? Department of Economics and Finance Working Papers. Available at: <http://ideas.repec.org/p/gua/wpaper/em200901.html>
- Barro, R.J. and J. Lee. 2001. International data on educational attainment: Updates and implications. *Oxford Economic Papers* 53(3): 541-563.
- Bledsoe, C. and M. Robey. 1986. Arabic literacy and secrecy among the Mende in Sierra Leone. *Man New Series* 21(2): 202-226.
- Easton, P. and M. Peach. 1997. The practical applications of Koranic learning in West Africa. Non-formal Education Working Group, Research Studies Series no.8. London: ADEA. Available at: http://pdf.usaid.gov/pdf_docs/PNACJ812.pdf
- El Sammani, M.O., Hassoun, I., Abdalla, B., and Hatim A. Gadir. 1985. 'Koranic' schools in Sudan as a resource for UPEL: Results of a study on Khalwas in RAHAD agricultural project. UNESCO document. Paris: UNESCO. Available at: http://unesdoc.unesco.org/Ulis/cgi-bin/ulis.pl?catno=64885&set=49F41391_2_52&gp=0&lin=1
- K.C., S., Barakat B., Goujon, A., Skirbekk, V., Sanderson, W., and W. Lutz. 2010. Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. *Demographic Research* 22 (15): 383 – 472.
- K.C. S., Barakat B., Goujon, A. Skirbekk, V., and W. Lutz. 2008. Projection of populations by level of educational attainment, age and sex for 120 countries for 2005-2050. IIASA Interim Report IR-08-038. Laxenburg: IIASA.
- Lutz, W., Butz, B., and S. KC. 2013 (forthcoming). *Oxford Handbook on World Population and Human Capital in the 21st Century*. Oxford University Press.
- Lutz, W., Goujon, A., K.C., S., and W. Sanderson. 2007. Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970-2000. *Vienna Yearbook of Population Research* 2007: 193-235.
- Riosmena, F., Prommer, I., Goujon, A., and S. KC. 2008. An evaluation of the IIASA/VID education-specific back projections. IIASA Interim Report IR-08-019. IIASA: Laxenburg.
- UN. 2007. *Principles and Recommendations for Population and Housing Censuses, Revision 2*. New York: United Nations Department of Economic and Social Affairs Statistics Division.
- UNESCO. 2006 [1997]. *International Standard Classification of Education ISCED 1997 [re-edition]*. Paris: United Nations Educational, Scientific and Cultural Organization.
- UNESCO. 2011. *Revision of the International Standard Classification of Education ISCED 1997 (ISCED)*. Paris: United Nations Educational, Scientific and Cultural Organization.
- UNESCO-UIS. 2011. *Global Education Digest 2011. Comparing Education Statistics Across the World*. Montreal: UNESCO Institute for Statistics.

Appendix Tables

Appendix Table 1: Documentation of data sources of the WIC 2012 dataset and adjustments

Country	Year	Data type	Data source	Data adjustments
Albania	2002	LSMS	World Bank	
Algeria	2002	PAPFAM	PAPFAM	
Argentina	2001	census	CELADE	7
Armenia	2001	census	IPUMS	2e
Aruba	2010	census	NSO	2e
Australia	2006	census	NSO	2e, 3e
Austria	2008	register	NSO	1, 2, 3e
Azerbaijan	2006	DHS	Macro Int.	6, 7
Bahamas	2000	census	Caricom	2e
Bahrain	2001	census	NSO	
Bangladesh	2004	DHS	Macro Int.	
Belarus	1999	census	IPUMS	
Belgium	2001	census	NSO (VUB)	2
Belize	2000	census	CELADE	
Benin	2006	DHS	Macro Int.	
Bhutan	2005	census	NSO	
Bolivia	2001	census	IPUMS	
Bosnia and Herzegovina	2010	LFS	NSO	
Brazil	2010	census	NSO	7
Bulgaria	2001	census	NSO	
Burkina Faso	2006	census	NSO	
Burundi	2010	DHS	Macro Int.	
Cambodia	2008	census	IPUMS	7
Cameroon	2004	DHS	Macro Int.	
Canada	2001	census	NSO	1e, 6
Cape Verde	2000	census	NSO	
Central African Republic	1995	DHS	Macro Int.	
Chad	2004	DHS	Macro Int.	
Chile	2002	census	CELADE	
China	2005	NPSS	NSO	2
Colombia	2005	census	CELADE	
Comoros	1996	DHS	Macro Int.	
Costa Rica	2000	census	CELADE	
Côte d'Ivoire	2005	DHS	Macro Int.	
Croatia	2001	census	NSO	
Cuba	2002	census	IPUMS	
Cyprus	2001	census	Eurostat	

Note: (1) category not available: no education; (2) category not available: incomplete primary; (3) category not available: primary ; (4) category not available: lower secondary; (5) category not available: upper secondary; (1e) category estimated: no education; (2e) category estimated: incomplete primary; (3e) category estimated: primary; (4e) category estimated: lower secondary; (5e) category estimated: upper secondary; (6) 5-year age groups interpolated; (7) cohort adjustments due to changes in educational system;

Country	Year	Data type	Data source	Data adjustments
Czech Republic	2001	census	Eurostat	2e
Dem. Rep. of the Congo	2007	DHS	Macro Int.	
Denmark	2001	census	Eurostat	1, 2
Dominican Republic	2002	census	CELADE	
East Timor	2009	DHS	Macro Int.	
Ecuador	2001	census	CELADE	
Egypt	2006	census	NSO	
El Salvador	2007	census	CELADE	
Equatorial Guinea	2000	MICS	UNICEF	2e, 4e
Estonia	2000	census	NSO	
Ethiopia	2011	DHS	Macro Int.	
Finland	2009	register	NSO	1, 2, 3e, 4e
France	2008	census	NSO	2
French Guiana	2008	census	NSO	2
French Polynesia	2007	census	NSO	2e
Gabon	2000	DHS	Macro Int.	
Gambia	2000	MICS	UNICEF	
Georgia	2002	census	NSO	
Germany	2010	LFS	NSO	1e, 2
Ghana	2000	census	IPUMS	7
Greece	2001	census	Eurostat	
Guadeloupe	2008	census	NSO	2e
Guatemala	2002	census	CCP	
Guinea	1996	census	IPUMS	
Guinea-Bissau	2000	MICS	UNICEF	4e
Guyana	2002	census	NSO	2e, 4e
Haiti	2005	DHS	Macro Int.	
Honduras	2001	census	CELADE	
Hong Kong	2006	census	NSO	
Hungary	2001	census	IPUMS	
Iceland	2010	LFS	NSO	1, 2, 6
India	2001	census	NSO	
Indonesia	2010	census	NSO	6
Iran	2006	census	IPUMS	7
Iraq	1997	census	IPUMS	2e
Ireland	2002	census	Eurostat	2
Israel	2004	LFS	UIS	
Italy	2001	census	Eurostat	
Jamaica	2001	census	Caricom	2e, 4e
Japan	2010	census	NSO	2e, 3e

Note: (1) category not available: no education; (2) category not available: incomplete primary; (3) category not available: primary ; (4) category not available: lower secondary; (5) category not available: upper secondary; (1e) category estimated: no education; (2e) category estimated: incomplete primary; (3e) category estimated: primary; (4e) category estimated: lower secondary; (5e) category estimated: upper secondary; (6) 5-year age groups interpolated; (7) cohort adjustments due to changes in educational system;

Country	Year	Data type	Data source	Data adjustments
Jordan	2004	census	IPUMS	2e, 4e, 7
Kazakhstan	2009	census	NSO	2e
Kenya	1999	census	IPUMS	
Kuwait	2005	census	NSO	
Kyrgyzstan	1999	census	IPUMS	
Laos	2005	FRHS	UNFPA	2e
Latvia	2000	census	Eurostat	2e
Lebanon	2007	census	NSO	6
Lesotho	2009	DHS	Macro Int.	
Liberia	2007	DHS	Macro Int.	
Lithuania	2001	census	NSO	
Luxembourg	2001	census	Eurostat	1e, 2
Macao	2006	census	NSO	
Macedonia	2008	LFS	NSO	4e
Madagascar	2008	DHS	Macro Int.	
Malawi	2008	census	IPUMS	
Malaysia	2000	census	NSO	2e
Maldives	2006	census	NSO	
Mali	1998	census	IPUMS	
Malta	2010	LFS	NSO	2e
Martinique	2008	census	NSO	2e
Mauritius	2000	census	NSO	
Mexico	2010	census	NSO	
Moldova	2004	census	NSO	
Mongolia	2000	census	IPUMS	2e
Montenegro	2003	census	NSO	
Morocco	2004	census	IPUMS	7
Mozambique	2007	census	NSO	
Myanmar	2007	FRHS	UNFPA	
Namibia	2007	DHS	Macro Int.	
Nepal	2001	census	IPUMS	
Netherlands	2001	census	Eurostat	1e, 2
Netherlands Antilles	2001	census	NSO	6
New Caledonia	2009	census	NSO	2e
New Zealand	2001	census	NSO	1e, 2e, 3e
Nicaragua	2005	census	CELADE	
Niger	2006	DHS	Macro Int.	
Nigeria	2008	DHS	Macro Int.	
Norway	2010	register	NSO	1, 2, 3e
Pakistan	1998	census	IPUMS	

Note: (1) category not available: no education; (2) category not available: incomplete primary; (3) category not available: primary ; (4) category not available: lower secondary; (5) category not available: upper secondary; (1e) category estimated: no education; (2e) category estimated: incomplete primary; (3e) category estimated: primary; (4e) category estimated: lower secondary; (5e) category estimated: upper secondary; (6) 5-year age groups interpolated; (7) cohort adjustments due to changes in educational system;

Country	Year	Data type	Data source	Data adjustments
Palestine	2007	census	IPUMS	
Panama	2010	census	CELADE	
Paraguay	2002	census	CELADE	
Peru	2007	census	IPUMS	
Philippines	2000	census	IPUMS	
Poland	2002	census	Eurostat	2e, 3e
Portugal	2001	census	IPUMS	
Puerto Rico	2000	census	IPUMS	
Qatar	2010	census	NSO	6
Republic of Congo	2005	DHS	Macro Int.	
Réunion	2008	census	NSO	2e
Romania	2002	census	Eurostat	2e
Russia	2002	census	NSO	
Rwanda	2002	census	IPUMS	
Saint Lucia	2001	census	Caricom	4e
Saint Vincent & Grenadines	2001	census	Caricom	4e
Samoa	2001	census	NSO	
Sao Tome and Principe	2009	DHS	Macro Int.	
Saudi Arabia	2004	census	NSO	
Senegal	2002	census	IPUMS	
Serbia	2002	census	NSO	
Sierra Leone	2004	census	IPUMS	
Singapore	2010	census	NSO	2e
Slovakia	2001	census	NSO	2e, 3e
Slovenia	2002	census	NSO	
Somalia	2006	MICS	UNICEF	4e
South Africa	2007	Community survey	NSO	
South Korea	2010	census	NSO	
Spain	2001	census	Eurostat	
Sudan	2008	census	NSO	
Suriname	2004	census	Caricom	2e
Swaziland	2006	DHS	Macro Int.	
Sweden	2010	register	NSO	1, 2
Switzerland	2000	census	NSO	2
Syria	2004	census	NSO	
Tajikistan	2009	LSMS	World Bank	
Tanzania	2002	census	IPUMS	
Thailand	2000	census	NSO	6
Tonga	2006	census	NSO	2e

Note: (1) category not available: no education; (2) category not available: incomplete primary; (3) category not available: primary ; (4) category not available: lower secondary; (5) category not available: upper secondary; (1e) category estimated: no education; (2e) category estimated: incomplete primary; (3e) category estimated: primary; (4e) category estimated: lower secondary; (5e) category estimated: upper secondary; (6) 5-year age groups interpolated; (7) cohort adjustments due to changes in educational system;

Country	Year	Data type	Data source	Data adjustments
Trinidad and Tobago	2000	census	Caricom	2e, 4e
Tunisia	2010	NSPE	NSO	5e
Turkey	2000	census	Eurostat	
Turkmenistan	1995	census	UIS	
Uganda	2002	census	IPUMS	
Ukraine	2001	census	NSO	
United Arab Emirates	2005	census	NSO	
United Kingdom	2001	census	NSO	1e, 2
United States	2005	ACS	IPUMS	
Uruguay	2004	census	IPUMS	
Vanuatu	2009	census	NSO	
Venezuela	2001	census	IPUMS	
Vietnam	2009	census	IPUMS	
Zambia	2002	DHS	Macro Int.	
Zimbabwe	2005	DHS	Macro Int.	

Note: (1) category not available: no education; (2) category not available: incomplete primary; (3) category not available: primary ; (4) category not available: lower secondary; (5) category not available: upper secondary; (1e) category estimated: no education; (2e) category estimated: incomplete primary; (3e) category estimated: primary; (4e) category estimated: lower secondary; (5e) category estimated: upper secondary; (6) 5-year age groups interpolated; (7) cohort adjustments due to changes in educational system;

Appendix Table 2: Proportion of population 25+ by sex and educational attainment

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
Albania	2002	Men	4.2	0.8	11.0	36.8	35.9	11.2
		Women	10.4	0.9	13.1	39.3	30.0	6.5
Algeria	2002	Men	27.1	6.0	15.3	26.4	18.5	6.8
		Women	52.4	4.5	10.7	15.2	13.0	4.3
Argentina	2001	Men	4.2	16.3	33.8	13.3	23.4	9.2
		Women	4.7	16.6	32.6	10.2	23.1	12.8
Armenia	2001	Men	1.4	0.7	5.9	10.3	60.2	21.5
		Women	2.8	1.6	6.1	9.1	60.9	19.5
Aruba	2010	Men	7.5	7.2	21.8	31.4	7.7	24.4
		Women	9.1	8.1	23.7	26.9	9.6	22.5
Australia	2006	Men	0.8	1.0	11.8	13.1	44.5	28.8
		Women	1.0	1.7	17.2	18.6	29.3	32.2
Austria	2008	Men	0	0	2.3	15.6	55.6	26.5
		Women	0	0	3.5	30.7	44.7	21.1
Azerbaijan	2006	Men	1.3	0.9	4.4	11.2	65.0	17.2
		Women	4.0	2.4	8.6	15.0	60.7	9.4
Bahamas	2000	Men	1.9	10.1	22.4	47.8	6.2	11.6
		Women	1.5	10.8	19.0	47.0	9.3	12.4
Bahrain	2001	Men	14.8	13.0	11.6	16.5	26.2	17.8
		Women	21.7	12.8	8.7	11.1	26.2	19.4
Bangladesh	2004	Men	40.6	15.3	16.0	11.4	6.7	10.1
		Women	54.2	16.7	14.6	7.2	3.6	3.7
Belarus	1999	Men	0.2	0.4	11.6	10.1	60.7	17.0
		Women	0.6	1.2	18.8	9.1	54.4	15.8
Belgium	2001	Men	4.4	0	16.8	25.2	26.6	27.0
		Women	5.1	0	21.6	24.7	23.5	25.0
Belize	2000	Men	11.3	32.3	34.5	10.6	1.5	9.7
		Women	10.9	33.8	34.0	10.6	1.9	8.8
Benin	2006	Men	49.1	23.1	13.5	7.1	4.0	3.1
		Women	77.6	11.9	6.4	2.6	1.0	0.5
Bhutan	2005	Men	54.6	22.5	1.0	12.5	2.6	6.8
		Women	81.8	8.6	0.5	5.9	0.9	2.3
Bolivia	2001	Men	15.1	24.6	18.5	14.8	13.5	13.5
		Women	36.0	22.8	14.2	10.3	10.5	6.1
Bosnia and Herzegovina	2010	Men	3.5	2.6	9.0	13.6	60.7	10.5
		Women	14.2	5.5	14.1	18.7	39.1	8.4
Brazil	2010	Men	11.0	18.0	21.6	15.3	24.2	10.0
		Women	10.8	16.6	20.3	14.7	25.1	12.5
Bulgaria	2001	Men	1.2	0.9	6.9	29.8	44.8	16.5
		Women	2.4	1.5	11.1	27.1	38.5	19.5
Burkina Faso	2006	Men	77.0	6.7	5.0	5.9	3.2	2.1
		Women	89.9	3.2	2.4	2.9	1.1	0.6
Burundi	2010	Men	47.4	24.9	19.8	3.4	2.1	2.3
		Women	77.5	20.6	15.3	2.6	1.5	1.1
Cambodia	2008	Men	18.4	26.0	28.0	16.1	8.6	3.0
		Women	37.3	27.9	20.6	10.1	3.3	0.9
Cameroon	2004	Men	22.0	18.2	32.7	12.3	9.4	5.4

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
		Women	41.7	19.9	25.6	7.0	4.0	1.9
Canada	2001	Men	1.3	0.8	8.8	8.3	29.6	51.4
		Women	1.6	0.9	9.2	8.3	32.7	47.3
Cape Verde	2000	Men	16.2	49.7	17.3	8.6	3.5	4.7
		Women	36.7	40.8	12.5	5.9	1.9	2.3
Central African Rep.	1995	Men	42.5	33.4	13.4	7.1	2.1	1.4
		Women	74.8	16.8	5.6	2.0	0.4	0.3
Chad	2004	Men	61.8	19.6	9.3	4.7	2.6	1.8
		Women	85.8	10.2	2.4	1.0	0.3	0.3
Chile	2002	Men	4.5	13.6	19.7	16.6	32.6	13.0
		Women	5.3	15.3	20.0	17.4	31.1	10.8
China	2005	Men	6.8	0.0	29.2	42.8	13.9	7.4
		Women	19.4	0.0	33.0	33.0	9.6	5.0
Colombia	2005	Men	9.8	20.7	29.5	5.9	18.5	15.6
		Women	9.5	20.2	29.1	6.2	18.7	16.3
Comoros	1996	Men	56.8	11.7	15.9	8.7	2.7	4.2
		Women	77.8	7.2	7.7	5.6	0.8	1.0
Costa Rica	2000	Men	6.8	20.4	37.9	8.7	11.3	15.0
		Women	6.4	21.7	36.8	8.6	12.0	14.5
Côte d'Ivoire	2005	Men	51.5	18.0	12.8	9.4	2.6	5.7
		Women	68.5	16.6	9.1	3.3	0.5	1.9
Croatia	2001	Men	1.3	3.2	9.8	15.7	54.8	15.2
		Women	5.0	6.9	16.3	21.4	37.6	12.8
Cuba	2002	Men	3.7	8.3	15.3	31.5	32.0	9.2
		Women	3.9	11.0	18.3	25.8	31.4	9.6
Cyprus	2001	Men	0.9	4.8	23.0	9.3	34.7	27.3
		Women	3.8	9.9	24.2	8.0	30.0	24.1
Czech Republic	2001	Men	0.4	0.0	0.2	10.7	72.6	16.0
		Women	0.5	0.0	0.3	26.7	60.0	12.4
Dem. Rep. of the Congo	2007	Men	9.8	20.4	13.6	30.0	19.2	7.0
		Women	36.5	27.2	11.6	16.7	6.5	1.5
Denmark	2001	Men	0.0	0.0	0.3	31.4	46.9	21.5
		Women	0.0	0.0	0.4	40.2	37.1	22.3
Dominican Republic	2002	Men	1.8	33.4	12.0	26.0	14.5	12.3
		Women	1.9	33.2	11.5	23.3	16.1	14.1
East Timor	2009	Men	39.0	17.4	11.0	7.2	19.8	5.6
		Women	56.7	11.4	10.6	7.2	11.7	2.3
Ecuador	2001	Men	10.8	20.8	30.0	11.2	11.0	16.2
		Women	14.7	21.5	26.1	11.0	12.0	14.7
Egypt	2006	Men	34.2	9.9	4.1	4.8	31.0	16.0
		Women	56.7	5.8	3.0	3.2	21.7	9.6
El Salvador	2007	Men	21.4	24.9	15.6	15.2	12.6	10.3
		Women	27.5	26.5	14.0	11.6	11.7	8.6
Equatorial Guinea	2000	Men	5.1	9.3	22.4	35.0	20.9	7.3
		Women	18.7	18.6	34.4	18.4	8.6	1.3
Estonia	2000	Men	0.2	0.3	8.3	19.6	44.8	26.7
		Women	0.2	0.6	11.4	15.5	38.3	33.9
Ethiopia	2011	Men	50.6	24.6	13.2	3.4	4.1	4.1
		Women	78.3	11.6	5.0	1.4	2.2	1.5
Finland	2009	Men	0.0	0.0	0.2	18.1	38.5	43.1

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
		Women	0.0	0.0	0.2	19.3	33.0	47.4
France	2008	Men	2.3	0.0	21.8	8.5	43.4	24.1
		Women	2.2	0.0	28.4	10.0	34.7	24.6
French Guiana	2008	Men	14.8	0.0	28.1	12.6	28.3	16.2
		Women	17.1	0.0	29.4	12.9	24.9	15.7
French Polynesia	2007	Men	5.5	6.3	20.3	19.1	31.6	17.2
		Women	5.2	5.7	17.0	20.8	33.2	18.1
Gabon	2000	Men	22.0	18.6	27.5	17.0	7.4	7.5
		Women	30.8	25.0	28.0	10.1	3.4	2.7
Gambia	2000	Men	48.5	20.1	5.5	14.7	6.5	4.8
		Women	69.1	14.3	5.5	6.7	2.3	2.0
Georgia	2002	Men	0.2	0.8	6.0	7.4	37.6	47.9
		Women	0.5	1.5	8.0	7.7	34.5	47.8
Germany	2010	Men	0.8	0.0	2.3	9.5	52.3	35.2
		Women	0.9	0.0	3.0	21.4	49.3	25.4
Ghana	2000	Men	41.2	3.7	10.0	23.6	14.2	7.3
		Women	59.9	4.3	8.7	15.5	7.4	4.3
Greece	2001	Men	2.3	6.4	33.6	9.7	30.8	17.1
		Women	6.8	10.0	34.9	7.3	27.7	13.3
Guadeloupe	2008	Men	2.3	10.6	26.0	13.1	32.7	15.2
		Women	1.7	10.9	24.6	13.9	31.8	17.0
Guatemala	2002	Men	29.8	31.0	18.4	7.6	7.4	5.8
		Women	44.3	25.0	13.7	6.0	7.3	3.6
Guinea	1996	Men	77.5	3.5	6.5	5.3	2.4	4.8
		Women	93.0	1.4	2.4	1.5	0.7	0.9
Guinea-Bissau	2000	Men	55.5	8.9	16.8	9.8	6.7	2.3
		Women	86.5	3.1	5.1	3.1	1.7	0.6
Guyana	2002	Men	3.0	9.2	23.1	28.2	28.8	7.8
		Women	3.0	7.6	18.1	27.9	34.1	9.3
Haiti	2005	Men	33.7	28.2	12.5	13.8	8.1	3.7
		Women	49.8	23.3	10.4	10.3	4.2	2.1
Honduras	2001	Men	26.6	29.9	24.7	5.7	7.4	5.7
		Women	28.1	28.3	23.8	5.4	9.9	4.5
Hong Kong	2006	Men	4.8	9.3	19.1	17.0	27.1	22.8
		Women	11.5	10.1	17.3	14.0	28.8	18.3
Hungary	2001	Men	0.7	0.6	7.2	29.3	48.5	13.8
		Women	0.9	0.8	14.4	34.7	37.4	11.7
Iceland	2010	Men	0.0	0.0	27.1	0.7	38.7	33.6
		Women	0.0	0.0	35.1	0.6	30.1	34.2
India	2001	Men	33.9	10.5	15.6	11.2	18.9	9.9
		Women	62.3	7.5	11.4	6.0	8.6	4.2
Indonesia	2010	Men	7.1	7.6	35.4	16.8	24.8	8.3
		Women	12.9	9.6	37.5	14.7	18.0	7.3
Iran	2006	Men	22.3	7.5	23.0	15.5	20.0	11.7
		Women	37.3	11.0	19.0	10.0	14.8	7.8
Iraq	1997	Men	30.8	8.3	24.7	10.5	10.4	15.3
		Women	52.9	10.7	17.4	5.6	5.6	7.8
Ireland	2002	Men	0.8	0.0	25.4	22.7	16.8	34.3
		Women	0.6	0.0	23.5	22.1	22.2	31.6
Israel	2004	Men	2.1	6.4	19.1	20.5	20.6	31.1

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
		Women	5.1	6.5	15.1	15.5	22.4	35.4
Italy	2001	Men	1.3	4.1	24.2	33.0	27.7	9.6
		Women	2.2	7.8	30.4	25.5	24.9	9.1
Jamaica	2001	Men	1.4	12.3	24.3	45.9	5.9	10.3
		Women	1.0	11.9	20.0	42.6	8.3	16.1
Japan	2010	Men	0.1	0.7	9.0	7.8	44.4	38.1
		Women	0.1	1.9	13.2	5.2	47.2	32.2
Jordan	2004	Men	14.2	6.5	17.7	13.8	21.2	26.7
		Women	24.5	6.2	15.4	10.9	19.6	23.4
Kazakhstan	2009	Men	0.2	1.0	2.0	10.6	63.7	22.3
		Women	0.4	1.8	3.7	10.1	58.9	25.1
Kenya	1999	Men	17.5	18.7	19.2	18.8	21.5	4.3
		Women	35.6	19.4	17.1	13.9	12.4	1.6
Kuwait	2005	Men	15.0	30.4	4.7	15.5	17.0	17.3
		Women	18.2	25.2	4.3	14.4	17.2	20.8
Kyrgyzstan	1999	Men	0.8	1.1	5.8	12.9	65.5	13.9
		Women	2.5	2.3	8.5	11.8	60.4	14.4
Laos	2005	Men	18.5	22.1	24.9	14.6	10.7	9.2
		Women	43.9	22.0	17.8	8.4	5.2	2.7
Latvia	2000	Men	0.7	0.2	6.2	27.9	39.9	25.0
		Women	1.4	0.5	8.8	24.1	36.0	29.2
Lebanon	2007	Men	9.3	7.4	25.9	27.1	14.4	16.0
		Women	17.3	5.8	21.0	29.7	13.7	12.7
Lesotho	2009	Men	22.2	35.9	21.6	7.4	7.4	5.6
		Women	5.2	37.4	34.5	10.0	7.3	5.7
Liberia	2007	Men	74.6	9.4	5.6	4.6	3.7	2.2
		Women	86.7	6.7	3.0	1.9	1.2	0.5
Lithuania	2001	Men	0.3	1.5	9.8	13.0	42.1	33.2
		Women	0.3	4.2	15.7	10.1	27.1	42.7
Luxembourg	2001	Men	6.7	0.0	21.2	15.8	30.0	26.3
		Women	8.6	0.0	30.0	20.4	25.4	15.6
Macao	2006	Men	3.3	10.1	23.3	26.1	22.5	14.7
		Women	8.3	11.3	21.8	24.6	20.2	13.9
Macedonia	2008	Men	1.6	8.1	8.5	20.2	47.9	13.8
		Women	6.5	17.0	10.1	22.0	33.1	11.2
Madagascar	2008	Men	20.9	43.0	20.0	9.5	3.7	3.0
		Women	28.3	41.2	19.0	7.0	2.6	1.9
Malawi	2008	Men	21.1	23.9	14.5	25.9	12.7	1.9
		Women	43.1	24.3	11.9	14.7	5.1	0.9
Malaysia	2000	Men	10.4	10.2	18.1	21.9	29.2	10.2
		Women	19.6	11.4	16.9	18.4	26.4	7.3
Maldives	2006	Men	28.3	26.3	29.6	9.8	1.3	4.7
		Women	26.6	31.8	28.2	9.8	0.6	3.0
Mali	1998	Men	80.5	7.1	4.9	2.7	3.4	1.3
		Women	91.1	3.9	2.5	1.2	1.1	0.3
Malta	2010	Men	0.7	2.2	22.6	45.4	6.3	22.8
		Women	0.6	7.3	27.1	43.3	6.0	15.7
Martinique	2008	Men	1.2	11.1	25.4	13.6	32.5	16.2
		Women	1.0	11.5	23.3	14.1	31.4	18.8
Mauritius	2000	Men	6.6	36.7	32.3	14.4	6.0	3.9

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
		Women	16.9	38.7	26.9	11.2	4.8	1.6
Mexico	2010	Men	7.8	15.5	21.3	26.2	13.6	15.5
		Women	10.7	16.3	22.0	26.5	11.8	12.8
Moldova	2004	Men	0.2	1.7	10.0	29.4	46.0	12.7
		Women	0.5	3.8	14.7	24.1	43.4	13.6
Mongolia	2000	Men	1.0	3.1	13.1	31.7	38.2	12.9
		Women	2.5	4.6	15.7	21.6	43.9	11.8
Montenegro	2003	Men	1.7	0.8	6.8	16.7	55.1	18.9
		Women	7.3	2.6	13.1	21.8	42.6	12.7
Morocco	2004	Men	47.4	8.3	19.5	10.5	7.7	6.6
		Women	69.9	5.0	9.9	6.7	4.6	3.8
Mozambique	2007	Men	62.6	18.0	5.6	8.4	4.5	0.9
		Women	83.8	8.4	2.2	3.4	1.8	0.4
Myanmar	2007	Men	14.4	6.5	37.1	22.8	12.1	7.1
		Women	12.0	12.1	44.5	14.5	8.4	8.5
Namibia	2007	Men	17.8	24.9	12.1	20.9	16.3	8.0
		Women	16.6	25.9	14.5	22.4	13.9	6.8
Nepal	2001	Men	54.4	6.0	10.0	6.5	17.7	5.4
		Women	83.5	2.6	4.4	2.5	6.2	0.9
Netherlands	2001	Men	3.3	0.0	11.0	20.1	38.1	27.3
		Women	4.1	0.0	16.8	27.9	31.4	19.8
Netherlands Antilles	2001	Men	0.7	11.3	27.3	35.1	14.6	11.0
		Women	1.0	12.8	30.8	33.7	13.4	8.2
New Caledonia	2009	Men	5.8	5.2	13.7	19.1	31.7	24.5
		Women	7.2	5.7	14.3	18.3	29.3	25.2
New Zealand	2001	Men	0.9	1.0	12.1	15.3	36.3	34.3
		Women	0.8	1.2	12.5	15.5	41.4	28.6
Nicaragua	2005	Men	26.4	26.5	20.5	7.6	10.0	9.1
		Women	27.8	25.0	20.1	7.3	11.4	8.5
Niger	2006	Men	80.2	9.6	5.5	2.1	1.0	1.5
		Women	89.9	6.1	2.6	0.8	0.4	0.3
Nigeria	2008	Men	29.6	4.7	20.6	6.3	23.8	15.0
		Women	49.4	6.4	17.9	4.7	13.2	8.3
Norway	2010	Men	0.0	0.0	0.3	22.9	47.4	29.5
		Women	0.0	0.0	0.4	25.1	40.9	33.6
Pakistan	1998	Men	53.1	6.2	11.2	10.5	13.9	5.2
		Women	79.2	3.3	6.0	3.7	5.7	2.1
Palestine	2007	Men	9.5	13.2	21.9	19.0	16.5	19.9
		Women	27.6	11.5	19.3	17.9	12.9	10.8
Panama	2010	Men	6.2	10.3	28.1	12.8	25.2	17.4
		Women	7.6	9.4	24.0	11.2	24.7	23.1
Paraguay	2002	Men	4.9	34.0	30.7	10.4	10.7	9.3
		Women	7.4	34.9	29.2	8.1	9.9	10.5
Peru	2007	Men	4.6	19.6	10.4	7.2	34.9	23.2
		Women	13.8	21.6	10.1	5.8	26.3	22.5
Philippines	2000	Men	3.3	18.1	27.8	3.3	23.6	23.9
		Women	4.0	16.5	29.4	3.1	22.2	24.7
Poland	2002	Men	1.3	0.7	0.5	21.0	62.6	14.0
		Women	3.1	0.8	0.3	27.3	50.5	18.0
Portugal	2001	Men	7.9	7.7	37.5	26.8	11.3	8.8

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
		Women	14.2	12.0	32.4	21.1	10.0	10.3
Puerto Rico	2000	Men	4.6	6.0	15.0	11.1	41.1	22.1
		Women	5.0	6.4	14.4	9.3	37.0	28.0
Qatar	2010	Men	3.8	25.7	23.3	11.3	21.1	14.9
		Women	6.1	19.6	11.3	9.8	21.6	31.6
Republic of Congo	2005	Men	8.4	17.0	33.8	21.2	9.7	9.9
		Women	27.4	18.4	34.8	13.0	4.2	2.2
Réunion	2008	Men	6.3	8.9	27.7	12.3	29.3	15.5
		Women	6.3	10.7	29.4	13.6	24.8	15.2
Romania	2002	Men	2.6	1.4	11.3	23.7	46.4	14.5
		Women	6.3	2.1	18.2	28.7	33.9	10.8
Russia	2002	Men	0.3	0.2	6.7	9.6	64.3	19.0
		Women	0.5	0.4	10.6	11.0	58.5	19.0
Rwanda	2002	Men	36.7	35.0	20.6	3.7	3.0	1.1
		Women	51.9	28.9	15.3	2.0	1.7	0.2
Saint Lucia	2001	Men	6.6	5.5	58.1	7.5	9.5	12.8
		Women	5.7	4.8	53.6	9.6	13.1	13.2
Saint Vincent & Grenadines	2001	Men	1.0	7.6	63.0	7.9	10.0	10.5
		Women	1.0	6.7	56.7	10.1	13.8	11.8
Samoa	2001	Men	0.8	2.4	49.8	28.1	4.9	14.0
		Women	0.7	1.9	47.5	32.2	6.1	11.6
Sao Tome and Principe	2009	Men	6.2	48.5	26.6	10.4	6.5	1.8
		Women	19.6	50.0	20.5	6.7	2.3	0.8
Saudi Arabia	2004	Men	12.3	6.1	19.2	18.2	19.2	24.9
		Women	35.7	8.7	12.8	10.4	12.7	19.6
Senegal	2002	Men	66.5	4.6	14.0	6.4	4.3	4.3
		Women	79.6	4.0	9.8	3.5	1.7	1.4
Serbia	2002	Men	0.7	1.2	13.8	20.7	48.6	14.9
		Women	7.0	6.7	20.1	20.0	34.8	11.3
Sierra Leone	2004	Men	60.0	8.2	8.5	9.9	8.4	5.0
		Women	78.7	6.8	4.9	4.4	3.6	1.7
Singapore	2010	Men	3.9	8.4	7.1	10.9	17.7	51.9
		Women	9.9	9.5	7.6	10.3	20.4	42.2
Slovakia	2001	Men	0.3	0.0	0.4	16.5	68.2	14.6
		Women	0.4	0.0	0.4	31.6	56.7	10.9
Slovenia	2002	Men	0.5	1.2	5.1	15.3	62.8	15.2
		Women	1.1	2.2	2.1	31.1	47.6	16.0
Somalia	2006	Men	48.2	4.5	13.3	11.5	16.8	5.6
		Women	79.4	3.4	7.2	5.6	3.7	0.7
South Africa	2007	Men	9.2	18.0	13.6	26.5	26.8	5.9
		Women	13.0	17.6	14.0	26.2	24.5	4.7
South Korea	2010	Men	1.8	0.6	8.4	9.5	38.6	41.1
		Women	7.4	1.4	14.2	10.9	35.9	30.4
Spain	2001	Men	1.8	13.5	24.2	30.6	15.5	14.4
		Women	4.1	16.3	25.6	27.8	12.1	14.1
Sudan	2008	Men	60.9	9.3	6.2	5.7	10.7	7.2
		Women	75.5	6.2	4.0	3.3	6.6	4.4
Suriname	2004	Men	0.4	7.1	32.9	39.7	13.6	6.4
		Women	0.8	13.1	30.6	35.1	14.8	5.6

Country	Year	Sex	None	Incomplete Primary	Primary	Lower Secondary	Upper Secondary	Post-secondary
Swaziland	2006	Men	18.7	21.3	20.0	11.4	17.0	11.6
		Women	20.8	23.6	22.8	12.6	11.9	8.3
Sweden	2010	Men	0.0	0.0	11.1	11.0	45.9	32.0
		Women	0.0	0.0	10.8	8.7	42.7	37.9
Switzerland	2000	Men	0.0	0.0	2.8	18.8	49.2	29.2
		Women	0.0	0.0	3.8	33.1	50.3	12.7
Syria	2004	Men	16.2	34.6	16.8	10.9	8.9	12.6
		Women	36.4	29.3	12.3	7.9	5.8	8.2
Tajikistan	2009	Men	1.9	0.0	2.8	11.4	64.2	19.7
		Women	4.1	0.8	7.7	18.2	62.0	7.2
Tanzania	2002	Men	22.7	18.3	49.8	6.1	1.8	1.2
		Women	44.7	13.6	37.1	3.3	0.6	0.6
Thailand	2000	Men	5.9	48.0	18.6	9.0	8.3	10.2
		Women	11.1	51.4	16.7	5.6	5.5	9.6
Tonga	2006	Men	1.3	1.4	26.0	46.2	11.3	13.7
		Women	1.5	1.4	29.2	46.2	11.1	10.6
Trinidad and Tobago	2000	Men	2.3	10.8	34.0	34.8	12.2	6.1
		Women	3.7	10.8	33.5	32.4	14.3	5.4
Tunisia	2010	Men	17.4	2.2	33.9	22.4	12.5	11.6
		Women	40.6	0.3	27.7	13.9	8.1	9.5
Turkey	2000	Men	6.8	5.1	50.3	11.3	16.2	10.3
		Women	26.4	7.4	45.9	5.4	9.3	5.5
Turkmenistan	1995	Men	0.8	1.1	4.6	12.7	65.0	15.9
		Women	2.4	2.6	7.7	15.4	62.4	9.6
Uganda	2002	Men	21.3	37.8	23.5	11.3	2.2	3.9
		Women	46.9	32.1	14.0	4.9	0.6	1.5
Ukraine	2001	Men	0.2	1.0	10.4	9.8	61.6	17.1
		Women	0.5	3.7	13.0	11.9	55.4	15.5
United Arab Emirates	2005	Men	11.3	15.5	13.1	17.1	27.3	15.7
		Women	10.9	10.8	7.5	11.0	34.8	25.1
United Kingdom	2001	Men	1.2	0.0	35.7	32.6	6.9	23.6
		Women	1.2	0.0	40.0	32.4	5.9	20.5
United States	2005	Men	0.9	0.9	4.7	7.3	51.1	35.2
		Women	0.9	0.8	4.5	7.2	52.9	33.7
Uruguay	2004	Men	1.6	14.3	37.7	26.5	10.0	9.9
		Women	2.0	14.0	36.6	21.8	12.0	13.6
Vanuatu	2009	Men	16.7	21.7	29.7	15.0	12.1	4.7
		Women	22.7	22.6	29.9	13.4	8.5	2.9
Venezuela	2001	Men	9.5	15.8	30.7	12.3	15.8	15.9
		Women	10.8	14.3	28.7	11.7	16.2	18.4
Vietnam	2009	Men	4.0	14.2	30.3	32.1	11.2	8.2
		Women	8.3	20.4	28.6	27.7	8.6	6.3
Zambia	2002	Men	10.8	25.0	25.6	20.6	11.3	6.6
		Women	28.0	34.8	19.8	11.4	2.8	3.1
Zimbabwe	2005	Men	6.8	17.2	19.8	13.0	36.2	7.0
		Women	16.3	24.7	20.8	12.2	22.7	3.2