ECONOMIC AND SOCIAL COMMISSION FOR WESTERN ASIA (ESCWA)

A STUDY OF AGE REPORTING IN SELECTED ARAB CENSUSES OF POPULATION

United Nations

Distr.
GENERAL
E/ESCWA/SD/2013/5
4 October 2013
ORIGINAL: ENGLISH

ECONOMIC AND SOCIAL COMMISSION FOR WESTERN ASIA (ESCWA)

A STUDY OF AGE REPORTING IN SELECTED ARAB CENSUSES OF POPULATION

United Nations
New York, 2013

13-0242

Acknowledgements

To the best of our knowledge, *A Study of Age Reporting in Selected Arab Censuses of Population* provides the first comparative assessment of age reporting and patterns of digit preference at both the national and subnational levels in Arab censuses of population. It was prepared by the Statistics Division of the Economic and Social Commission for Western Asia (ESCWA) under the supervision of Mr. Marwan Khawaja and Mr. Ismail Lubbad. The contributions of Mr. Evan Brand, Mr. Raffi Shirinian and Ms. Zeina Sinno of the Statistics Division of ESCWA are acknowledged. It is hoped that findings of this study will be useful to national statistical offices as well as to researchers, students, journalists and the public at large.

We express our gratitude to the Integrated Public Use Microdata Series International (IPUMS-I), a project of the Minnesota Population Center at the University of Minnesota, in the United States for providing microdata files for Egypt, Iraq, Jordan, Morocco and the Sudan. We express our gratitude to the Palestinian Central Bureau of Statistics for providing micro census data for Palestine, and to national statistical offices for providing data in response to our questionnaire.

CONTENTS

Page

Acknowledgements iii
Abbreviations and acronyms vii
Introduction 1

Chapter

I. OVERVIEW OF CENSUS-TAKING IN THE ARAB REGION 2

A. Historical background on Arab censuses before World War II 2
B. Censuses in Arab countries after World War II 3
C. Census-taking methods 4

II. DATA AND METHODS 5

A. Description of microdata files 5
B. Variables used in the analysis 5
C. Methods 8

III. DESCRIPTIVE ANALYSIS OF AGE-SEX REPORTING AND AGE HEAPING IN RECENT POPULATION CENSUSES 12

IV. ANALYSIS OF AGE HEAPING AT THE SUBNATIONAL LEVEL IN RECENT POPULATION CENSUSES 22

V. CONCLUSIONS AND RECOMMENDATIONS 27

References 38

LIST OF TABLES

1. Description of census samples 5
2. Definitions of literacy in census files 6
3. Definitions of urban areas in census files 7
4. Summary statistics for predictors 23
5. Pairwise correlations weighted by district counts between the Whipple index and predictors 23
6. Effects of district characteristics on the Whipple index 25

LIST OF FIGURES

1. Dates of censuses in 19 Arab countries 3
2. Scatterplots of four predictors vs. Whipple index for the 2008 census of the Sudan 11
3. Population pyramids of 12 Arab countries, various years 12

CONTENTS (*continued*)

Page

4. Birth cohorts by sex from multiple censuses in eight Arab countries 14
5. Sex ratio by birth cohort in 12 Arab countries, various years 16
6. Country-level Whipple indices for 12 Arab countries by sex, various years 18
7. Myers index for 10 Arab countries, most recent census 19
8. Bachi index for 10 Arab countries, most recent census 20
9. United Nations age-sex accuracy index for 13 Arab countries, most recent census 21
10. Boxplots of district-level Whipple indices 22
11. Scatterplots of per cent literate vs. wealth index in the 1982 census of Morocco and the 2008 census of the Sudan 24
12. Scatterplots of nonlinear relationships in the 1996 and 2006 censuses of Egypt 25
13. Scatterplot of per cent female vs. Whipple index in the 2008 census of the Sudan 26

ANNEXES

I. National data 28
II. Subnational data 33

ABBREVIATIONS AND ACRONYMS

CAPI	Computer assisted personal interview
GIS	Geographical information systems
GPS	Global positioning systems
IPUMS-I	Integrated Public Use Microdata Series International
ESCWA	Economic and Social Commission for Western Asia
NIU	Not in universe
PAPI	Paper and pencil interview
PDA	Personal digit assistant
PES	Post enumeration surveys
UNACI	United Nations age-sex accuracy index

Introduction

The first population characteristic given in census results is the age and sex distribution. Estimates of fertility, migration, morbidity and mortality rates all depend on age and sex data. Such data are also important for use in population projections, and for assessing the age- and sex-specific socioeconomic and health conditions of a population.

Unfortunately, data on age and sex often suffer from reporting errors and irregularities, which are sometimes serious enough to impact the usability of the data. One of the common errors is age misstatement, which is referred to as a content (or response) error. There are a number of reasons for age misstatement, but it most commonly results from the respondent simply not knowing his or her exact age or the age of the household members for whom he or she is reporting.

When a large number of ages are estimates, age distribution contains visible peaks at preferred ages. Often these are ages that end in zero or five, but they can also be other preferred digits that are of cultural significance. This pattern is known as "age heaping", and can be easily measured. Age heaping causes artificial distortion of the age structure of a population, resulting in inaccurately reported population indicators, and leads to erroneous conclusions about demographic patterns in the population.

It is therefore necessary to assess the reliability of age and sex data before use in any demographic analysis or projection. Indeed, an assessment of age and sex data is one of the most important steps in the evaluation of population censuses.

The purpose of this report is to do the following: (1) provide an overview of census taking in the Arab region; and (2) assess the accuracy of age-sex reporting and the patterns of digit preference in recent censuses of population in selected Arab countries at the national and subnational levels.

Two kinds of data are used in this study. Special tabulations of age and sex data were obtained directly from national statistical offices. In addition, the study analyses sample microdata files from population censuses in six countries: Egypt, Iraq, Jordan, Morocco, Palestine and the Sudan.

The first chapter of the study presents an overview of census taking in the Arab region and the second chapter explains the data and methods used for analysis in the report. The third chapter assesses the accuracy of age-sex reporting and patterns of digit preference in recent population censuses in some Arab countries at the national level. It presents graphical analyses of age-sex structures to examine data consistency. A number of common indices are used to assess the quality of age-sex reporting, including the United Nations age-sex accuracy index (UNACI) and the Whipple, Myers and Bachi indices. Due to limited availability of single-age data, this report will assess the severity of age heaping at the national level for 11 Arab countries, and at the subnational level for six Arab countries. Trends of age misstatement over time are only possible for countries where data from multiple census years are available.

The fourth chapter uses national census data from selected Arab countries to provide a subnational analysis of age heaping in population censuses in the region and examines the association between age heaping and selected household characteristics. The latter is accomplished by building a linear model that uses the Whipple index as the response variable and known or suspected correlates of age heaping as explanatory variables, namely age, literacy, sex, area of residence and household wealth. Since income data are not available, a wealth index is constructed instead from a linear combination of asset ownership variables, closely following the procedure used by the Demographic and Health Surveys.[1]

The objective of building a linear model is to quantify the amount of age heaping attributable to characteristics of the population, rather than census-taking issues, in six Arab countries.

---

[1] Rutstein and Johnson, 2004.

I. OVERVIEW OF CENSUS-TAKING IN THE ARAB REGION

Broadly, two principal eras of census-taking can be distinguished in the Arab region: before and after World War II. In the majority of Arab countries, the data from the first censuses of the nineteenth and early twentieth centuries are of variable quality. Counting methods were rather sketchy, which resulted in significant under-enumeration of populations.[2]

A. HISTORICAL BACKGROUND ON ARAB CENSUSES BEFORE WORLD WAR II

During the nineteenth century, enumerations of Arab populations were mainly undertaken by the Ottomans. During the first half of the twentieth century, censuses were generally partial or incomplete, carried out under the British, French, Spanish or Italian mandates.

Egypt is one of the first countries to have conducted a census. Evidence from ancient Egypt, including papyrus manuscripts, monuments and temples confirm that in 3340 B.C., a census was carried out. More recently, the Ottomans conducted enumerations, beginning in the sixteenth century (1570-1590), in the Fertile Crescent (modern-day Iraq, Jordan, Lebanon, Palestine and the Syrian Arab Republic). The Ottoman censuses continued up to 1881-1893,[3] and expanded to include Egypt. Nevertheless, these censuses were incomplete and unreliable. Censuses were also conducted under the French Mandate in Algeria throughout the nineteenth century, but these censuses, too, are considered unreliable.[4]

Changing objectives of censuses: From fiscal and military to socioeconomic goals

Early census enumerations were primarily done for fiscal and military reasons. For example, once the Ottoman conquered a territory, they conducted censuses to take stock of material and human resources. In Egypt, "from the sixteenth century, tax and land records contained lists of households".[5]

The 1848 Egyptian census was unique in that it not only served to count homes, but also to describe each individual within the household. This census marked the birth of statistical practice in Egypt, breaking with the tradition of administrative fiscal counts. In 1868, the successors of Muhammad Ali organized a second census and created a statistical office in 1870. This change would also happen in the Maghreb countries, but not until the twentieth century. However, compared to western countries, modern population censuses have had a very short history in most of the Arab region.

Partial or incomplete censuses, or general enumerations, were conducted in most Arab countries during the first half of the twentieth century. Egypt conducted a census in 1907, and subsequent censuses were conducted every ten years. In other Arab countries, most censuses conducted during the first half of the twentieth century were under the mandates of Britain, France, Italy or Spain. For example, under the British Mandate, two censuses were conducted in Palestine, in 1922 and 1931. Under the French Mandate, censuses were conducted in Lebanon (1921, 1932) and the Syrian Arab Republic (1922, 1927). These censuses are considered incomplete by many researchers. Also considered incomplete are the censuses and enumerations covering parts of Morocco (the "Ex-Zone Nord", Tarfaya and Ifni), conducted in 1930, 1935, 1940 and 1950 under the Spanish Mandate. No more reliable were the censuses carried out in the "Ex-Zone Sud" part of Morocco in 1921, 1926, 1931, 1936 and 1951/52 under the French Mandate.

In Algeria eight administrative censuses were taken from 1901 to 1948; in Libya censuses were taken in 1931 and 1936; and in Mauritania one census was taken in 1944. In Tunisia general enumerations were carried out in 1906, 1911, 1921, 1926, 1931 and 1936.

---

[2] Courbage and Fargues, 1975.

[3] The censuses of 1897, 1906 and 1914 concerned only some *wilayat.* See Courbage and Fargues, 1996, p. 147.

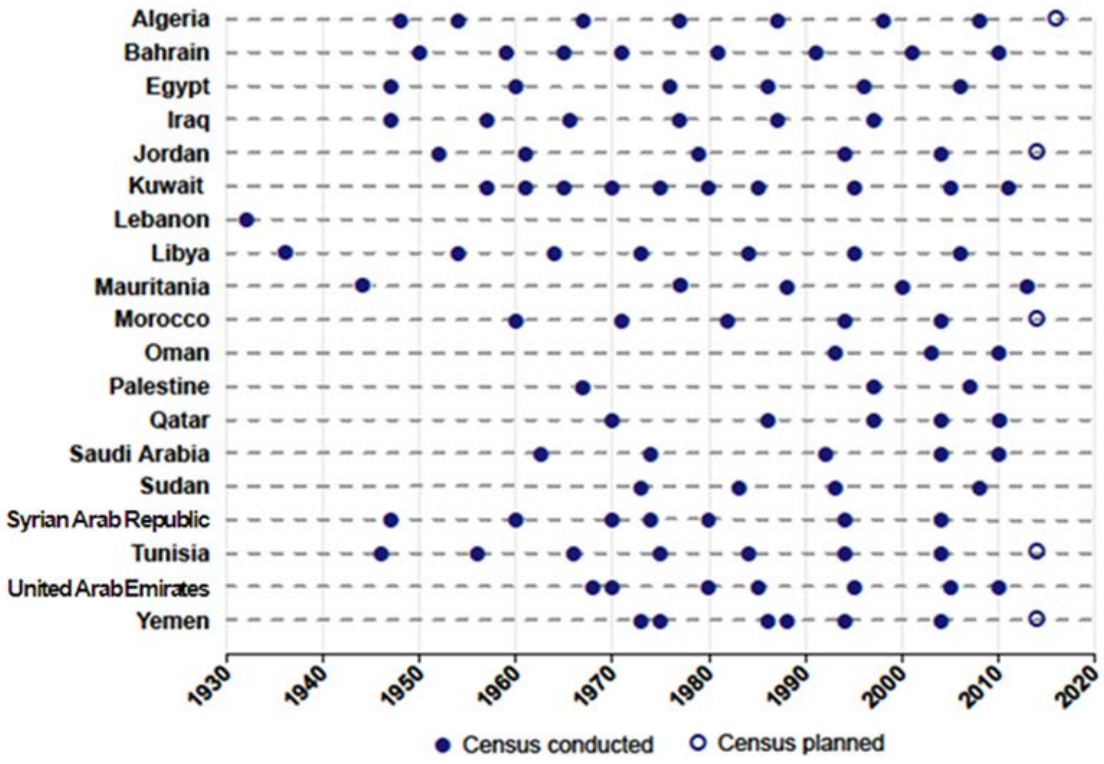
[4] Ibid., p. 310. Beginning in 1851, the French conducted population count approximately every five years (23 counts).

[5] Alleaume and Fargues, 1998.

B. Censuses in Arab countries after World War II

All Arab countries except Lebanon have carried out one or more modern censuses from the 1950s until 2013 (figure 1). Qatar and Yemen did not undertake a modern census until the 1970s, and Oman conducted its first census in the 1990s. Algeria, Bahrain, Egypt, Iraq, Kuwait, Libya, the Syrian Arab Republic, Tunisia and the United Arab Emirates have carried out at least six modern censuses. Jordan, Mauritania, Morocco and Qatar have carried out five modern censuses. The Sudan and Yemen have carried out four modern censuses, while Oman, Palestine, and Saudi Arabia have carried out less than four modern censuses. The 1967 census in the West Bank and Gaza Strip was supervised by the Israeli Central Bureau of Statistics, but was conducted by the Israeli military.

Figure 1. Dates of censuses in 19 Arab countries



Source: Annex table AI.1.

Note: Some censuses were incomplete or covered only part of the present day population of the country. Some censuses have been omitted from the figure.

Prior to the 1960s, and for some countries even up to the 1970s, population censuses suffered from many limitations, which led to inaccurate counts and irregularities in age and sex data. For example, nomadic and seminomadic populations were often markedly underreported. Other segments of the population, such as young children and especially young girls, were also likely to be underreported.

In addition to these coverage errors, there were problems in age reporting. Age misreporting caused noticeable distortions in the age-sex distributions of many populations in the region. For low-income countries in particular, ages of respondents were often unknown. Even in the most recent censuses, systematic errors of age declaration are still found in the age-sex distributions of many Arab countries: ages that end in 0 or 5 are often overreported, especially among women. This difference between males and females is typically wider in rural areas than it is in urban areas.

C. CENSUS-TAKING METHODS

A brief questionnaire on census-taking methodologies and the quality of the most recent census was sent by the Economic and Social Commission for Western Asia (ESCWA) to each of its member countries.[6] Eleven countries answered the questionnaire. This section gives a brief summary of the responses, in order to shed some light on the technical aspects of data collection that relate to the quality of census data on population characteristics.[7]

The methodology used for the most recent census in all respondent countries was a classical enumeration, with direct interviews by a census enumerator and generally one questionnaire. Some countries used more than one questionnaire, depending on the type of the enumerated population: Qatar (seven questionnaires); Jordan (four questionnaires); Morocco (three questionnaires: one for the nomad population; one for parts of the population such as military persons; and one for ordinary households), and Yemen (two short questionnaires). The period of enumeration is highly variable, from one day in Iraq to 40 days in Kuwait and 53 days in Egypt.

The traditional paper and pencil interview (PAPI) is the most frequently used method of data collection. A new data collection method, however, has been used in Oman and the United Arab Emirates: the computer assisted personal interview (CAPI), which is carried out with a pocket or handheld computer known as a personal digit assistant (PDA). Kuwait, Oman and Qatar have used the Internet as a supplemental tool in data collection.

During the preparation stage, most countries made use of global positioning systems (GPS) to help locate addresses within the enumeration areas (with the exception of Jordan, Morocco and Palestine). Countries that used GPS also used geographical information systems (GIS), not only during the preparation stage, but also during the enumeration stage itself (Kuwait, Oman, Qatar and Yemen). Six countries used GIS for publishing census results: Egypt, Iraq, Oman, Qatar, the United Arab Emirates and Yemen.

Only Egypt, Iraq, Jordan, Morocco and Palestine reported that they carried out post enumeration surveys (PES) to evaluate coverage errors. However, according to reports by the national statistical offices, several Arab countries used demographic analysis to evaluate their last census (Egypt, Iraq, Oman, Palestine, Qatar, the Syrian Arab Republic and Yemen). Kuwait, Oman, Palestine, Qatar and the United Arab Emirates used administrative records to check census coverage.

Not all countries publish their census coverage rate. Of those that do, undercoverage rates range from around 3 per cent in Palestine to 8 per cent in Egypt.[8] Only two countries, Kuwait and the United Arab Emirates, reported non-response rate (or refusal rate), although both countries reported this rate as 0 per cent.

[6] The member countries of ESCWA are Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestine, Qatar, Saudi Arabia, the Sudan, the Syrian Arab Republic, Tunisia, the United Arab Emirates and Yemen.

[7] Details are available in annex I tables AI.3-AI.6.

[8] The undercoverage rate for the 2004 census of Tunisia was reported as 1 per cent by Mrabet, 2013.

II. DATA AND METHODS

Two sets of data are used in this report. The first is special tabulations on age-sex distributions obtained mainly from the national statistical offices of the 11 member countries that responded to the questionnaire. The second set of data comes from sample microdata files readily available from IPUMS-I, a project of the Minnesota Population Center in the United States.

A. DESCRIPTION OF MICRODATA FILES

Ten census microdata files from six countries are used in the regression analysis in this report: Egypt (1996, 2006), Iraq (1997), Jordan (2004), Morocco (1982, 1994, 2004), Palestine (1997, 2007) and the Sudan (2008). All files except those from Palestine were provided by IPUMS-I. The Palestine census files were obtained directly from the Palestine Central Bureau of Statistics.

All census files are samples, ranging from 5 to 15 per cent of the full census populations. Four of the census files make use of sample weights and the other six do not (table 1).

TABLE 1. DESCRIPTION OF CENSUS SAMPLES

Country	Year	Fraction (*Percentage*)	Weighted	Number of areas
Egypt	1996	10	Yes	278
	2006	10	No	316
Iraq	1997	10	No	44
Jordan	2004	10	No	44
Morocco	1982	5	No	65
	1994	5	No	60
	2004	5	No	60
Palestine	1997	10	Yes	62
	2007	10	Yes	76
Sudan	2008	15	Yes	202

The unit of analysis in the second part of this report is the designated subdistrict (governorate, region) available in the data files. In the Palestine census files, two area units are available: governorate and locality. Neither can be used as the unit of analysis on its own: there are too few governorates and many of the localities are too small. Instead, localities whose populations are greater than 10,000 are treated as separate observations, and the remaining localities in each governorate are pooled. In other words, the 62 observations in the Palestine 1997 file represent 46 populous localities, plus the 16 Palestine governorates.

B. VARIABLES USED IN THE ANALYSIS

The analysis uses a few known or suspected correlates of age heaping as explanatory variables: age, literacy, sex, area of residence and household wealth. A brief explanation of relevant variables used in the analysis follows.

Age

Two approaches are used to collect age data in censuses and surveys, expressed in completed solar years: completed age (age at the individual's last birthday), and/or the date of birth (year, month and day), which is the preferred approach. Demographers point to many sources of age reporting errors, including the following:

- Misunderstanding of the meaning of age, whether the most recent, the next or the nearest birthday. In particular, children under 1 year may be reported as 1 year of age;
- Rounding to attractive ages, particularly those ending in 0 or 5 (age heaping);
- Use of different calendars in the same country: Western, Islamic or lunar.

In the Egypt 1996 census, age data were collected by first asking for the date of birth, and if unknown, asking for the age in years. In Morocco, either age in years or date of birth was accepted, and in Palestine, both age in years and date of birth were asked. For the census in Egypt 2006, Jordan 2004 and the Sudan 2008, only age in years was asked. For the census in Iraq 1997, only date of birth was asked.

For the regression model, median age is chosen over mean age, in part because of the skewness of the age distribution, but also because many of the census files top-code ages over 98 or 99. This top-coding does not affect the median age, but biases the mean age downward slightly. It should be noted that weighted censuses sometimes result in district-level median ages that are neither integers nor end in 5.

Literacy

Literacy is reported for the population aged 15 years or over. The definitions used in the various censuses are fairly similar (table 2). The computation of this variable from the data files is straightforward, except for the censuses of Iraq 1997 and Jordan 2004. The literacy variables in those censuses contain many missing values and not in universe (NIU) observations. Of the age 15+ population, these values make up close to 10 per cent in Iraq and 17 per cent in Jordan.

In Iraq, literacy was only asked of individuals who do not hold an education certificate. In Jordan, literacy was only asked of individuals who have not completed elementary education and who are not currently enrolled in school. The literacy variable is recoded to literate, then, if an individual has either completed primary school or is currently enrolled in school. After these observations are recoded, literacy information is only missing from 1 per cent of the observations in the files of Iraq and Jordan. These observations are dropped when calculating district-level literacy rates.

TABLE 2. DEFINITIONS OF LITERACY IN CENSUS FILES

Country	Definition
Egypt	The 1996 census does not define literacy; it is constructed from individuals who answered they can read and write or that they have an education certificate. Literacy is derived from the education attainment question in 2006, and implies ability to read and write.
Iraq	Asked of individuals without an education certificate. Ability to read and ability to write are asked separately; only those who can both read and write are considered literate.
Jordan	Ability to read and write. Asked of individuals who have not completed elementary education and who are not currently enrolled in school.
Morocco	Ability to read and write a simple paragraph.
Palestine	Implies ability to read and write. Asked as part of education level question.
Sudan	Ability to read and write a simple sentence.

Source: IPUMS-I literacy variable documentation, https://international.ipums.org/international-action/variables/LIT#comparability_section.

Wealth index

In the absence of income data, a modified version of the wealth index procedure used by the Demographic and Health Surveys is calculated at the household level.[9] The procedure uses principal components analysis to construct a linear combination of asset ownership variables from the first component, known as the Filmer-Pritchett method.[10] This report deviates from the Filmer-Prichett method in its treatment of ordinal asset variables with more than two levels (such as no kitchen, shared kitchen and private

[9] Rutstein and Johnson, 2004.

[10] Filmer and Pritchett, 1998.

kitchen). Instead of creating dummy variables for each ordinal variable level, the levels are simply recoded as consecutive integers, starting at 1 for the level associated with the lowest wealth. Although this assumes equal distance between levels, it has been shown to be an improvement over using dummy variables, in part because it incorporates knowledge about the order of the levels.[11]

In cases where the order of asset variable levels is unclear, ambiguous levels are combined. For instance, the variable for sewage disposal sometimes includes levels for septic tank and public sewage network. Since it is not obvious which of the two is associated with higher wealth, the levels are combined, and the variable becomes a binary indicator of connection to any sewage disposal system.

In addition to asset variables, a crowding variable is included, equal to the negative of the number of persons over the number of rooms in the household. The negative is used so larger values correspond to higher levels of wealth.

If a household reports "unknown/missing" for an asset, it is assumed the household does not own the asset. Generally, missing values make up a relatively small portion of the asset variable (< 1 per cent). It should be noted that, since the wealth index is centred at 0 for each census and based on different asset variables for different countries, only within-census comparisons of wealth may be meaningful.

Per cent urban

The definition of urban areas varies from country to country, but is generally related to population size, socioeconomic conditions and/or administrative considerations (such as municipal boundaries) (table 3). In Palestine, refugee camps are considered neither urban nor rural. In this report, per cent urban, per cent rural and per cent in camps all add to 100 (rather than per cent urban and per cent rural adding to 100). Per cent urban and per cent in camps are therefore negatively correlated by design, but this avoids creating missing values for the per cent urban variable in areas that are exclusively made up of refugee camps.

TABLE 3. DEFINITIONS OF URBAN AREAS IN CENSUS FILES

Country	Definition
Egypt	Administratively defined at the village level. Villages within the governorates of Cairo, Alexandria, Port-Said and Suez are exclusively urban. Villages within the capital districts of other governorates are defined as urban, as are the capital villages of most of the non-capital districts.
Iraq	Administratively determined. Urban areas are located inside the boundaries of municipality councils (Al-Majlis Al-Baldei) or within the boundaries of the City of Baghdad. Other areas are considered rural.
Jordan	Localities that had populations of 5,000 or more in the 1994 census.
Morocco	Urban/rural is not available for Morocco.
Palestine	Any locality whose population amounts to 10,000 persons or more is considered urban, as well as all governorates/district centres regardless of their size. Localities whose populations range from 4,000 to 9,999 persons and having at least four of the following services are also included as urban: public electricity network, public water network, post office, health centre with a full-time physician, and a school offering a general secondary education certificate. Refugee camps are neither classified as urban nor rural.
Sudan	Localities of administrative and/or commercial importance or with 5,000 or more inhabitants. The nomad population is included in rural.

Source: IPUMS-I urban variable documentation, https://international.ipums.org/international-action/variables/URBAN#comparability_section.

[11] Kolenikov and Angeles, 2009.

C. METHODS

1. *Assessing the quality of age and sex data*

There are several approaches for assessing the quality of age and sex data in censuses. Graphical techniques, such as age-sex pyramids, permit one to perform a qualitative assessment by visually identifying response errors. Several indices that allow one to quantify the magnitude of particular patterns of age errors are also available, which operate under the assumption of a rectangular digit distribution. Software programs such as SINGAGE in the Population Analysis System (PAS), developed by the International Programs Center (IPC) of the United States Bureaus of the Census, can be used to calculate many of these indices.[12]

Using census data provided by member countries, data available in publications and data tabulated from microdata files (obtained from IPUMS-I), graphical representations of age-sex data will be provided and age accuracy indices (Whipple, Myers, Bachi and UNACI) will be calculated to assess age reporting errors that result from digit preference.

Population pyramids

The population distribution by age and sex takes the shape of a pyramid, unless mortality and fertility levels are irregular, or the population in question has experienced important levels of migration. When the population is displayed by single years of age, it may also show age misstatement, if the pyramid displays peaks at certain ages, most frequently ages ending with 0 and 5.

Graphical cohort analysis

Comparing data from two or more censuses shows if the age pattern of the population at two or more census dates is consistent. Assuming no significant international migration, the size of each cohort should decline in successive censuses, due to mortality. In addition, the curve of age structure for censuses should follow the same pattern, in the absence of census errors.

Sex ratios

The sex ratio measures the number of males per 100 females. At birth, this ratio should be between 102 and 107, because male births tend to slightly outnumber female births. The sex ratio is then expected to decline gradually with age, due to lower female mortality rates. Major deviations from a smooth declining pattern can be attributed to the following:

- Fluctuations in demographic characteristics, such as mortality and migration rates;
- Fluctuations in sex ratio at birth;
- Misreporting of ages;
- Different enumeration rates for males and females at different ages.

Whipple index

The Whipple index was developed to measure preference for or avoidance of terminal digits 0 and 5. The index is calculated as the number of individuals between the ages of 23 and 62 whose reported age ends in 0 or 5, over the expected number of individuals whose ages should end in 0 or 5 in the 23-62 age group, multiplied by 100:

[12] Arriaga, 1994.

$$WI = 100 \times \frac{\sum[(P]_{25} + P_{30} + \cdots + P_{60})}{\frac{1}{5} \times \sum[(P_{23}] + P_{24} + \cdots + P_{62})}$$

Where P_k stands for the number of individuals whose reported age is *k*.

It follows that a Whipple index score close to 100 indicates little or no age heaping in the data, and higher values indicate more severe heaping.

Whipple index scores can be grouped in the following categories:

- Highly accurate data Less than 105
- Fairly accurate data 105-109.9
- Approximate data 110-124.9
- Rough data 125-174.9
- Very rough data 175 and more

In practice, when data at the individual level are available, it is convenient to generate a dummy variable as follows:

$$WI_i = \begin{cases} 500 \times I_{\{a_i \bmod 5 = 0\}}, & 23 \leq a_i \leq 62 \\ missing, & otherwise \end{cases}$$

Where a_i is the age of individual *i*.

The interpretation of this dummy variable is the Whipple index calculated for each individual. This allows one to compute Whipple indices for different area units by simply cross tabulating the mean of the Whipple index variable (ignoring missing values) by the area unit of interest.

Myers index

The Myers index is conceptually similar to the Whipple index, except that it can detect heaping at ages ending in each digit from 0 to 9. The theoretical range of the Myers index is from 0 to 90, where 0 indicates no age heaping and 90 indicates the extreme case where all recorded ages end in the same digit. Terminal digits with positive values are overreported, while digits with negative scores are underreported, compared to expected digit frequencies.

Bachi index

The Bachi index, as an indicator of the general extent of heaping, is similar to the Myers index. It involves applying the Whipple method repeatedly to determine the extent of preference for each terminal digit. Like the Myers index, the Bachi index is equal to the sum of the positive deviations from 10 per cent. It considers the population between 23 and 72 years. The theoretical range of the Bachi index is 0 to 180, where 0 indicates no age heaping, and 180 indicates that a single terminal digit was reported for all ages. For each terminal digit, positive values represent a preference, while negative values represent avoidance.

United Nations age-sex accuracy index (UNACI)

UNACI is based on sex ratios and age ratios, both of which are calculated in five-year age groups for populations below age 70. The sex ratio refers to the number of males per 100 females in each age group. It is calculated by first summing up successive differences (regardless of sign) in sex ratios and then taking their mean. The age ratio is calculated as 100 times the number of persons in a given age group divided by

the averages of the two adjacent age groups. It is calculated by first summing up deviations of the ratios from 100 for each age group. Then, mean deviation is computed. The final UNACI score is obtained by weighting the sex ratio score by 3 and the age ratio score by 1.

Scores obtained from UNACI are interpreted as follows:

- Less than 20 Reasonably reliable data
- Between 20 and 40 Fairly unreliable data
- Over 40 Quite unreliable data

It should be noted, however, that the index does not take into consideration the normal decline of sex ratios by age, and distortions in the age structure due to large labour migration or forced migration caused by wars or natural disasters. This is especially important to keep in mind when using this index for the heavily labour importing countries in the Gulf.

2. *Building a model*

A separate linear statistical model is constructed for each census. Each model regresses the Whipple index in each subnational area (e.g., district) on known or suspected correlates of age heaping, and is weighted by the estimated population of the area. For most census files, the unit of analysis for the model is the smallest available geographical distinction for each country, typically the district. The covariates considered are per cent female, per cent literate, per cent of the population in urban areas, median age, wealth index and per cent in camps (Palestine only).

The form of the model is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

Where Y is the Whipple index, β indicates regression coefficients, X indicates predictors, k is the number of predictors considered for the census and ε is the error term, which is assumed to have constant variance.

It is possible to combine the data and construct a single multilevel model, with districts nested within censuses (or even districts nested within years nested within countries), but there would only be a small number of units at each level above the district. In cases where the number of two- or three-level units is small, the separate model approach is justifiable.[13] Furthermore, the variables common to all censuses would be limited, and the methods of calculation differ for some variables. Only per cent literate, per cent female and median age could reasonably be included as predictors in a combined model.

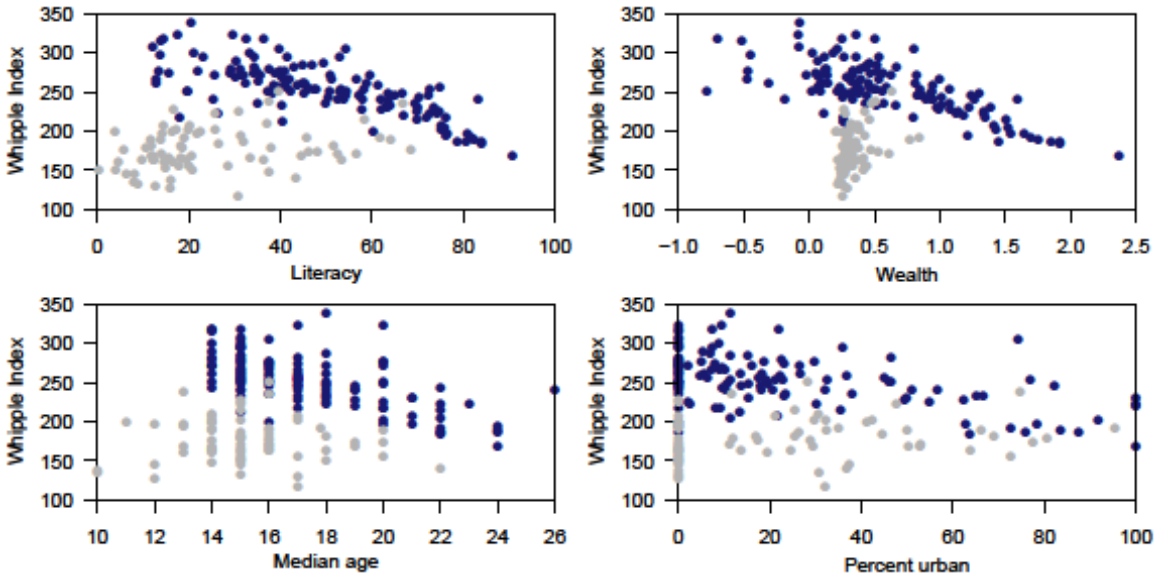
To address possible issues with multiple testing, after a model is built for each census, *p*-values from all models are pooled and adjusted using Holm's method. Holm's method is chosen because it controls for the familywise error rate and does not assume independence of the *p*-values.[14]

Data from South Sudan are considerably different from data from the rest of the Sudan (figure 2). The most recent census, taken in 2008, was divided into the Sudan and South Sudan, and the data are modelled separately. Likewise, IPUMS-I has recently split off data for South Sudan from the 2008 file, and data for both countries are available separately on its website.

[13] Goldstein 1999.

[14] Aickin and Gensler, 1996.

Figure 2. Scatterplots of four predictors vs. Whipple index for the 2008 census of the Sudan



Note: Blue points are counties in the Sudan and grey points are counties in South Sudan.

III. DESCRIPTIVE ANALYSIS OF AGE-SEX REPORTING AND AGE HEAPING IN RECENT POPULATION CENSUSES

The basic objectives of this chapter are to do the following:

- Examine the accuracy of age-sex data in selected Arab countries;
- Examine, when data are available, trends over time in the accuracy of age-sex data.

The analysis makes use of population pyramids and other graphical depictions of the data show areas where errors may be present. In addition, the data will be analysed using the Whipple, Myers, Bachi and UNACI indices.

Population pyramids

Single year population pyramids for selected Arab countries are shown in figure 3. These pyramids show data from the most recent censuses, and use the national populations when data are disaggregated by nationality/citizenship. Data from Egypt, the Sudan and Yemen are clearly affected by age heaping. Data from Morocco and the Syrian Arab Republic are also affected by heaping, albeit to a lesser extent. In the rest of the countries, age misreporting appears to be minor and concentrated at older ages. Of course, abnormal pyramids of Gulf countries, showing large numbers of working-age individuals, are a result of international labour migration.

Age heaping appears to be more severe for women than for men, and appears to be more extreme in older cohorts than in younger cohorts. Age heaping noticeably decreases over time in Morocco, but improvement is not as obvious in Egypt.

Figure 3. Population pyramids of 12 Arab countries, various years

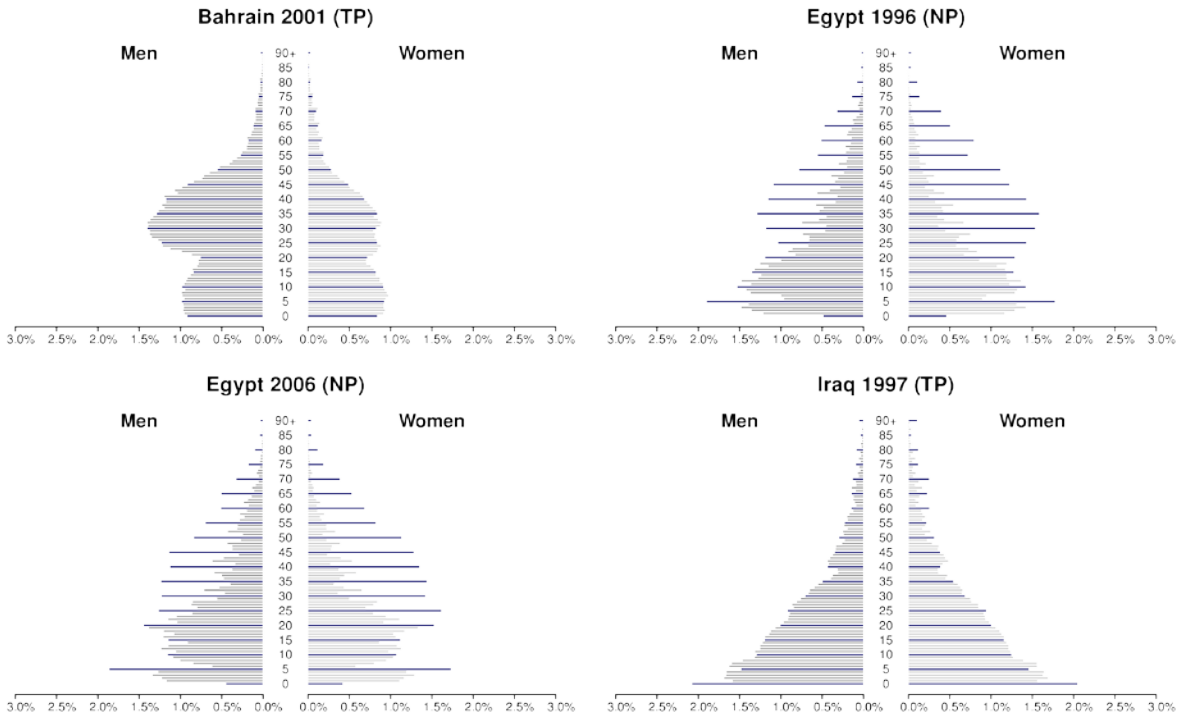


Figure 3 (*continued*)

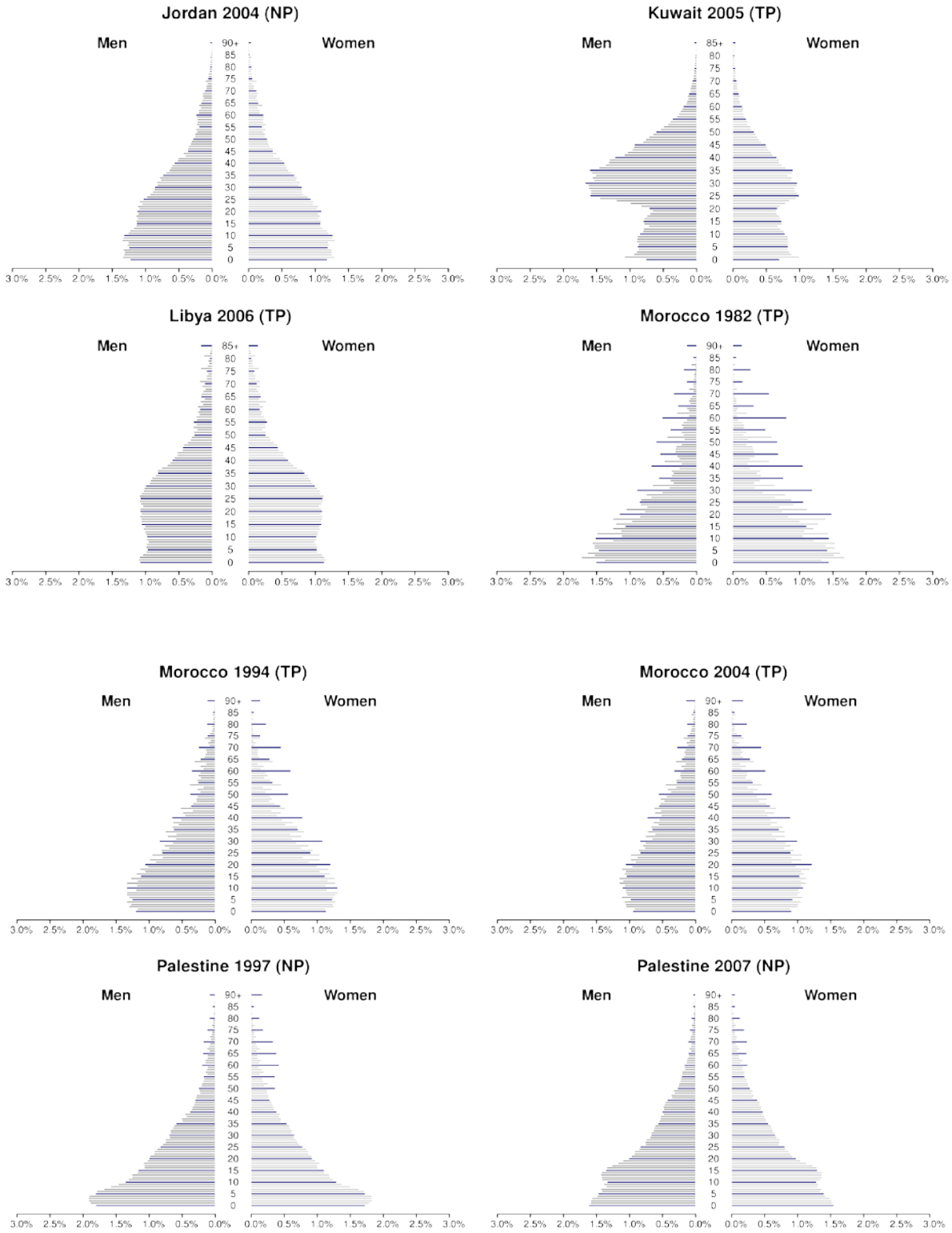
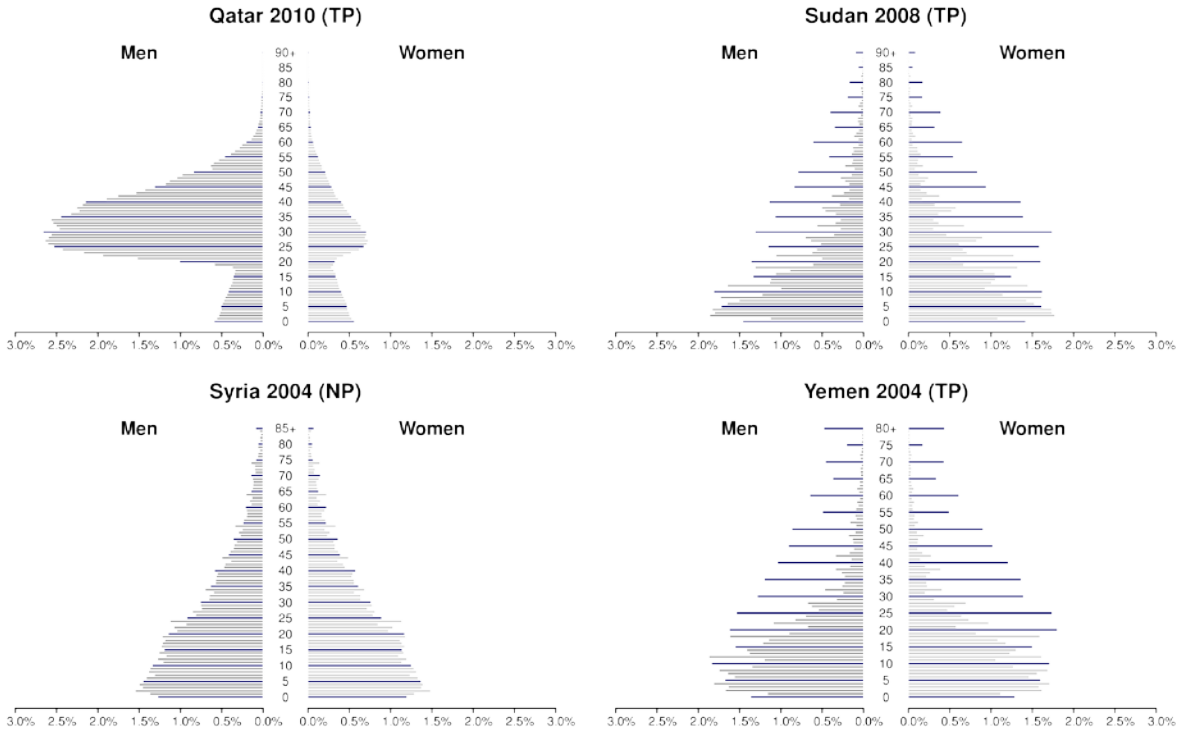


Figure 3 (*continued*)



Note: (NP) indicates national population; (TP) indicates total population.

Graphical cohort analysis

Figure 4 shows birth cohorts separately for each sex in Arab countries for which data from multiple census years are available. Use of total population in Kuwait and Qatar highlights the distorting effect of immigration when we follow five-year cohorts over time.

For the national population of Egypt, it is likely that children were under-enumerated in the 1996 census, as the 1986-1996 cohorts, individuals who were under 10 years of age at the time of the 1996 census, are significantly larger in the 2006 census. There could also have been an under-enumeration of children born between 1980 and 1990 based on the 1995 census of the national population of the United Arab Emirates compared to the size of that cohort as reported in the 2005 census. The size of the cohort in Morocco and Palestine appears to be fairly consistent between censuses.

Figure 4. Birth cohorts by sex from multiple censuses in eight Arab countries

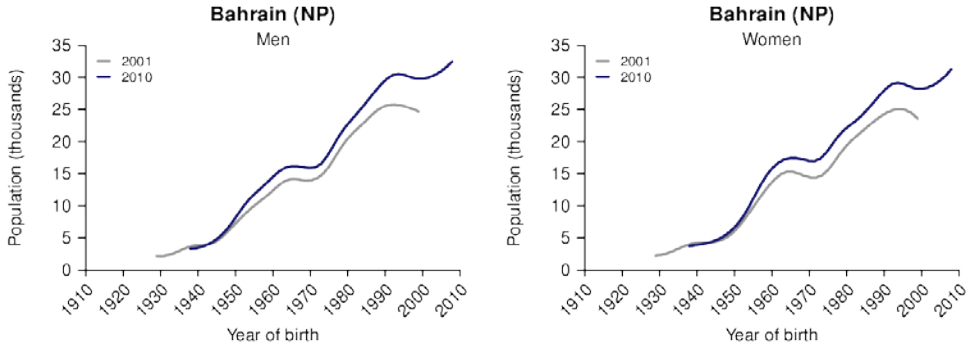


Figure 4 (*continued*)

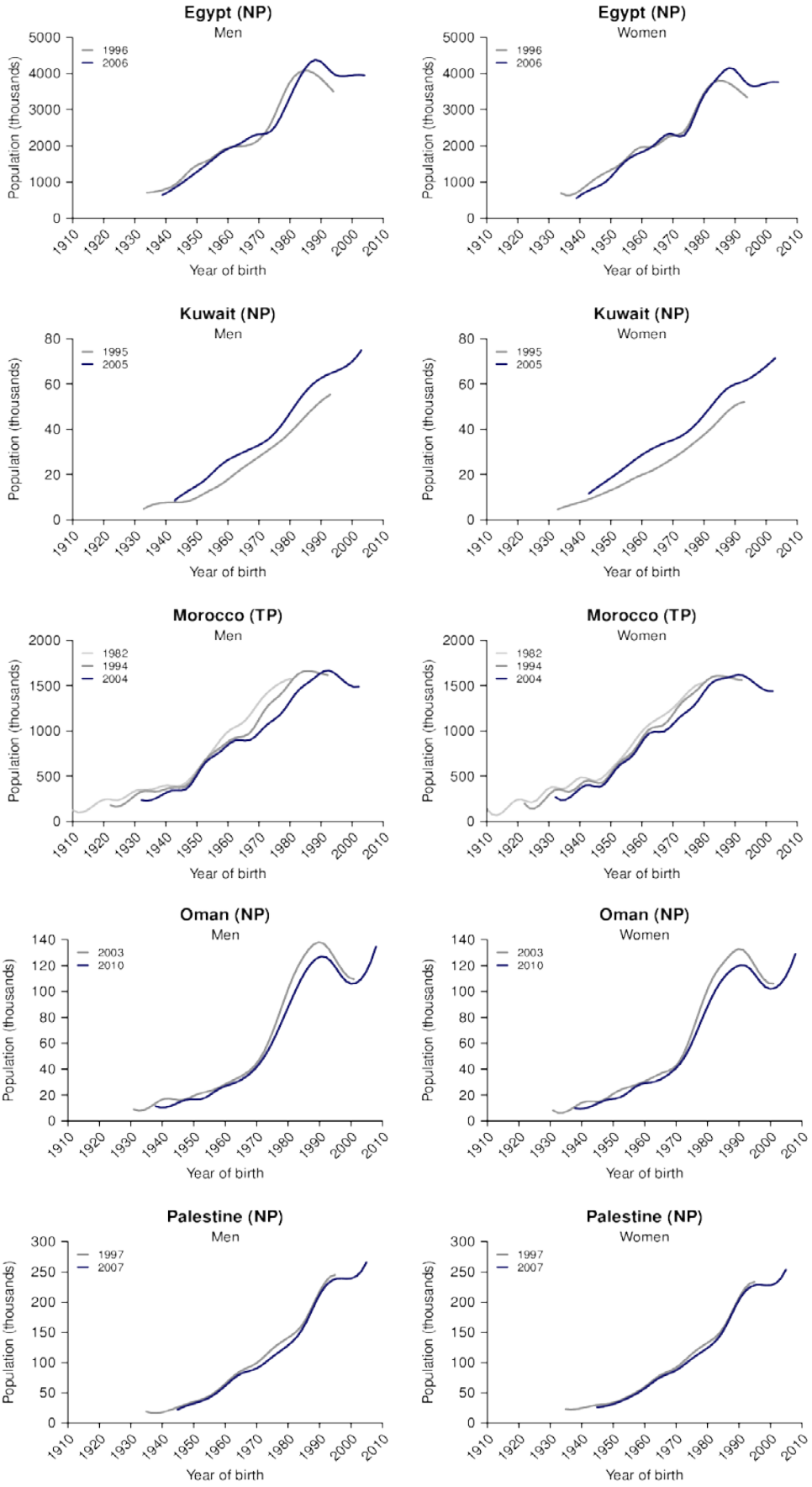
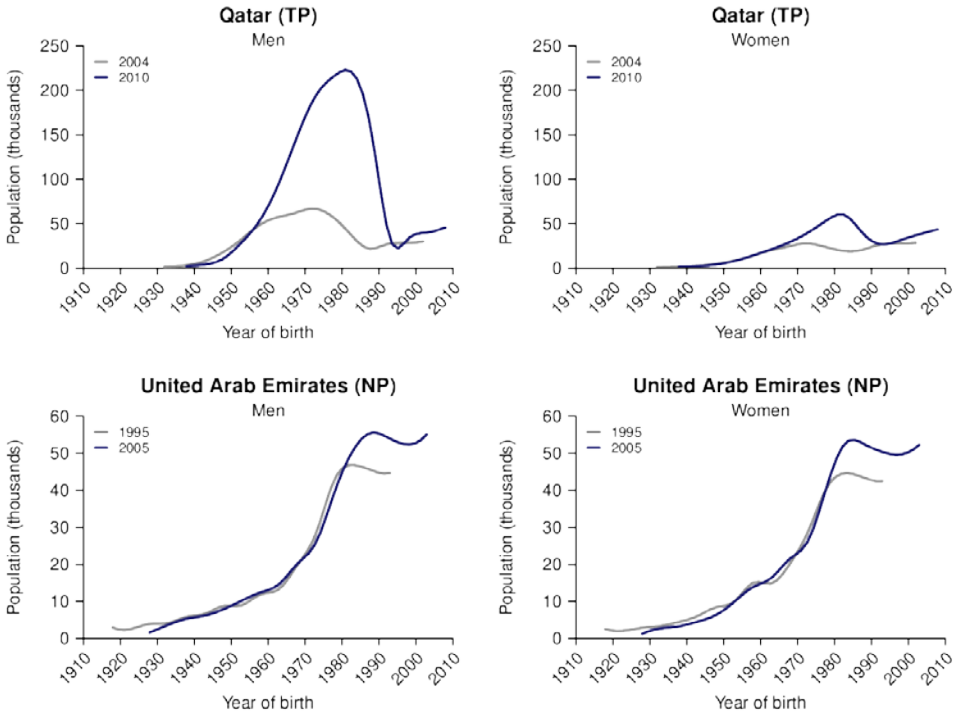


Figure 4 (*continued*)



Note: (NP) indicates national population; (TP) indicates total population.

Pattern of sex ratio by five age groups

Sex ratios by birth cohort are shown in figure 5. A smoothly increasing sex ratio by year of birth (the same as a smoothly decreasing sex ratio by age) is not seen in every country. Instead, sex ratios are higher than expected in older cohorts in some countries, suggesting possible under-enumeration of women or misreporting of their ages. In Egypt, for example, the sex ratio is close to 120 for those born before 1950, according to the 2006 census. The sex ratio is also greater than 100 in Yemen for those born before 1950, according to the 2004 census. It should be noted that international migration does not account for the high sex ratios of Bahrain and the United Arab Emirates for certain birth cohorts, since the graphs show national populations. For Qatar, however, total population is displayed, so international migration does explain the high sex ratios. In Libya, data from the 2006 census show slightly lower sex ratios than expected at young ages and increasing ratios thereafter. Volatile sex ratios at older ages in Libya suggest sex-selective age misreporting or coverage errors or both.

Figure 5. Sex ratio by birth cohort in 12 Arab countries, various years

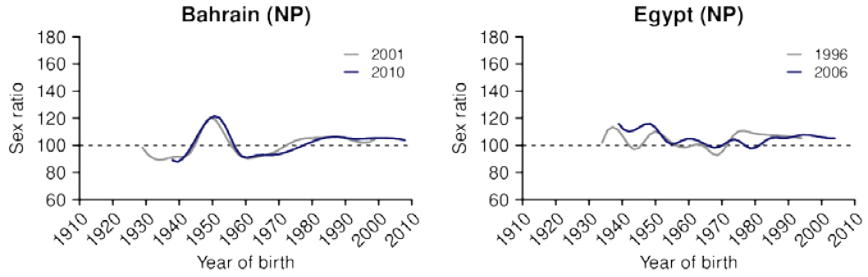
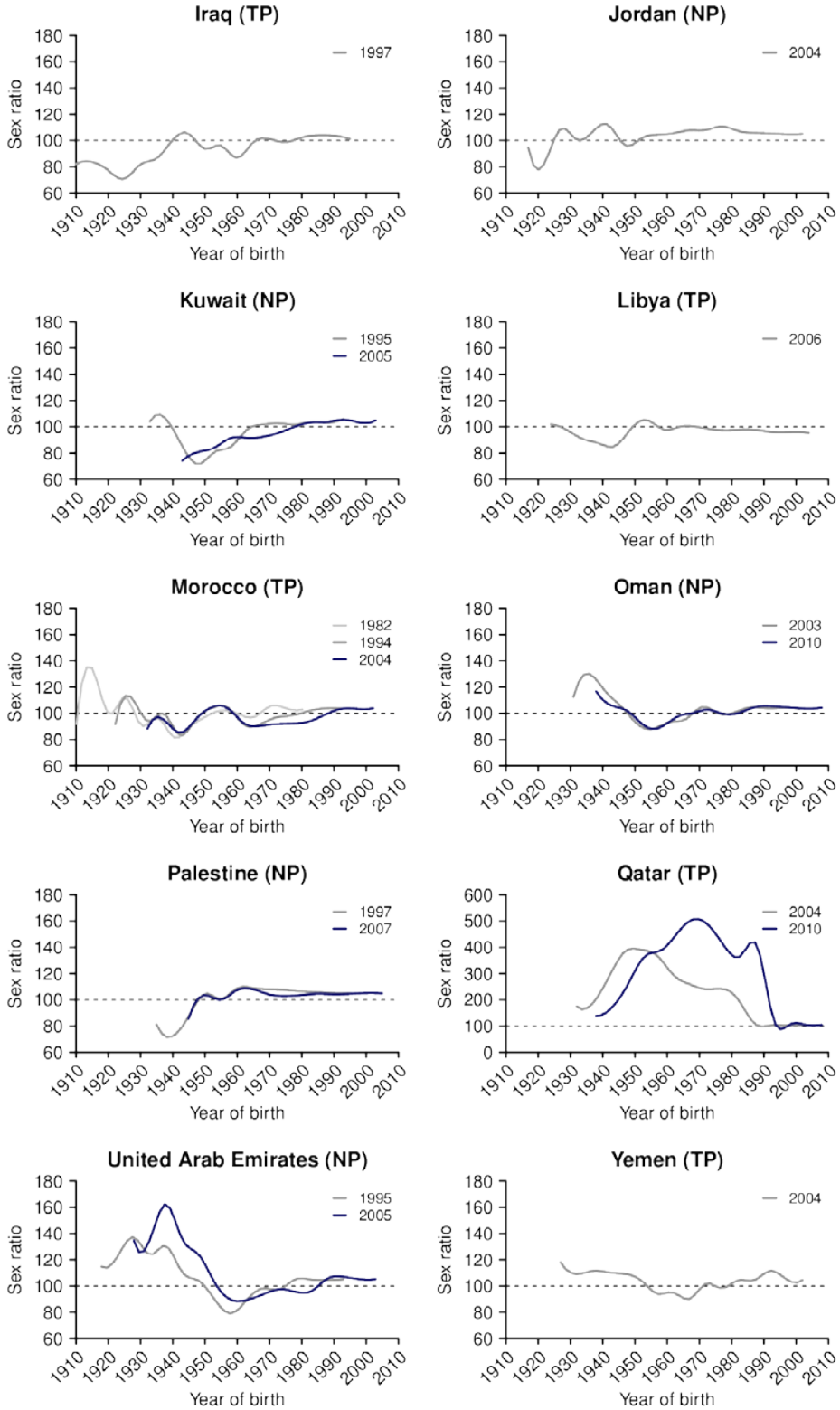


Figure 5 (*continued*)

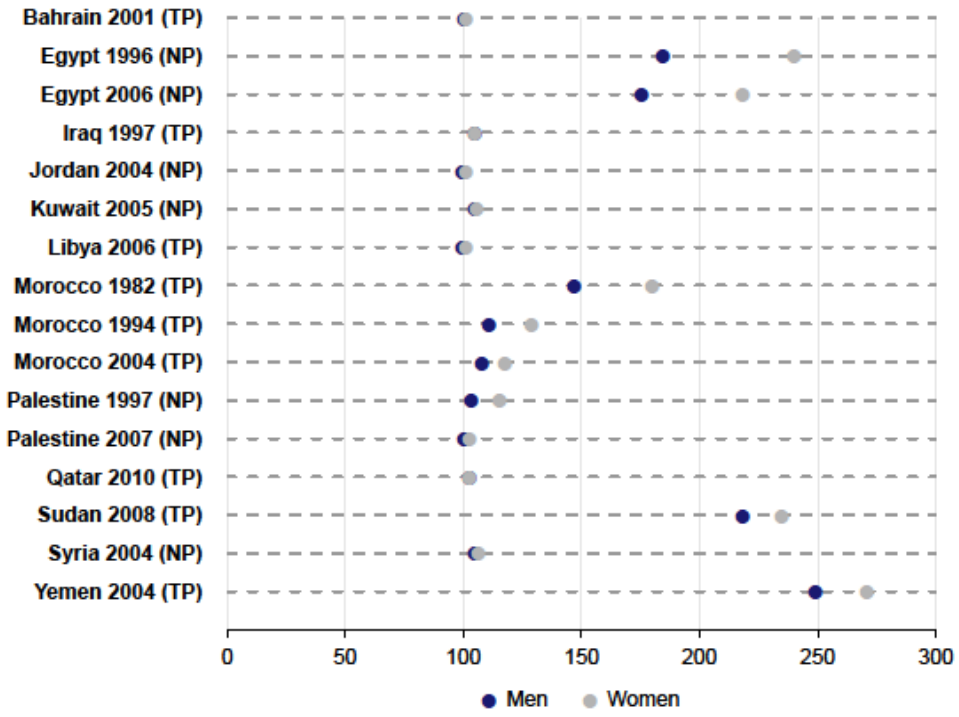


Note: (NP) indicates national population; (TP) indicates total population.

Whipple index

Figure 6 shows country-level Whipple indices for 12 Arab countries by sex. There appears to be little age heaping at terminal digits 0 and 5 in the most recent censuses of Jordan, Kuwait, Libya, Palestine, Qatar and the Syrian Arab Republic, since the index is close to 100. Data from multiple censuses are available for Egypt, Morocco and Palestine, and all show improvements in age data over time. Severe age heaping is present in the censuses of Egypt, the Sudan and Yemen, and also in the 1982 census of Morocco. Heaping at 0 and 5 is more extreme for women than it is for men in many censuses.

Figure 6. Country-level Whipple indices for 12 Arab countries by sex, various years



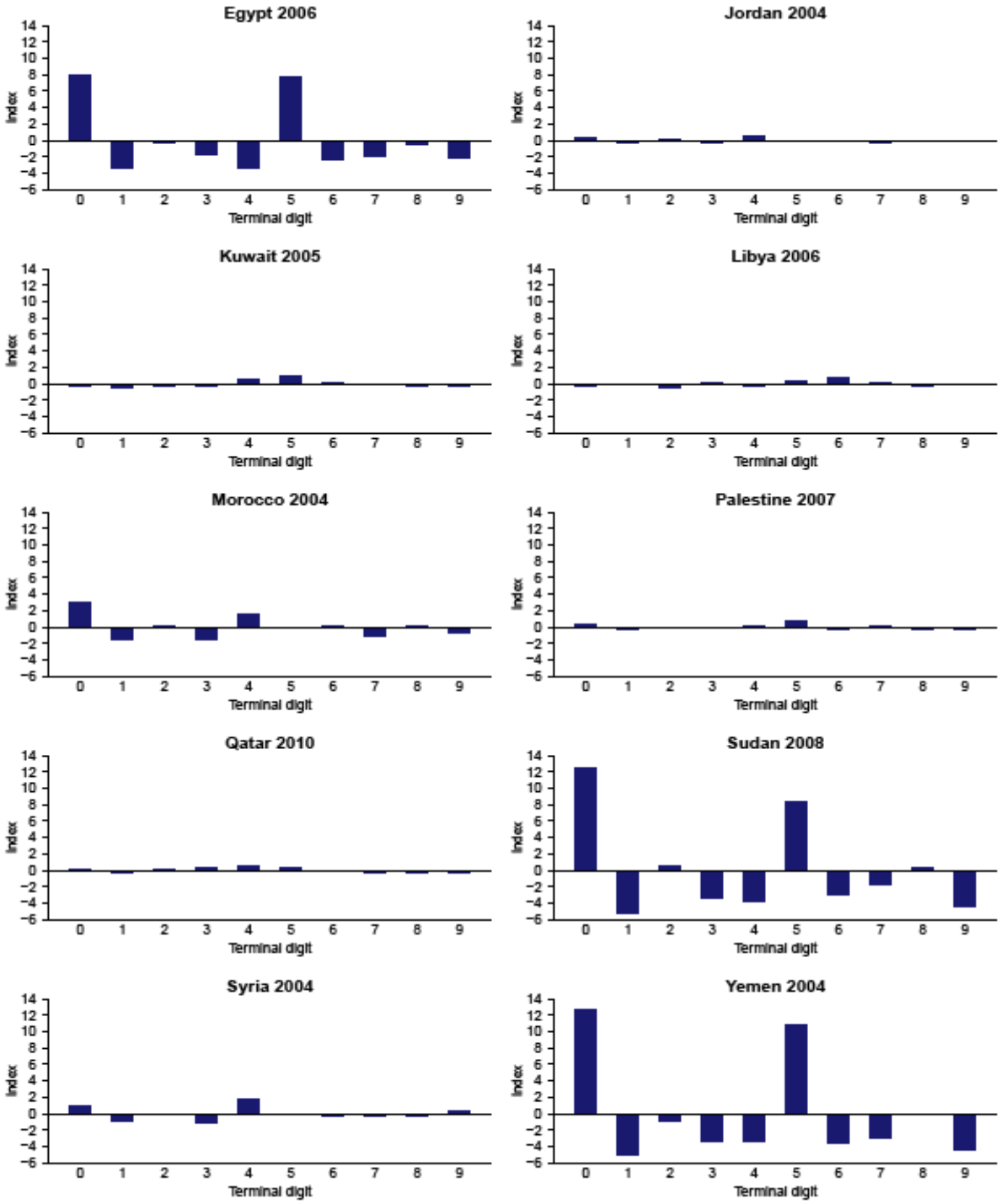
Note: (NP) indicates national population; (TP) indicates total population.

Myers index

Figure 7 shows the Myers index for 10 Arab countries, according to the most recent census. A Myers index score of less than 10 suggests little or no heaping. Jordan, Kuwait, Palestine, Qatar and the Syrian Arab Republic all scored less than 10. Higher scores suggest the age data are of poor quality. Egypt, the Sudan and Yemen each have a score greater than 30 on the Myers index.

In all countries, ages ending in 1, 3, 6, 7, 8 and 9 are underreported to some extent. Scores for digits 0 and 5 are highest in Egypt, the Sudan and Yemen, in agreement with the Whipple index, which is also highest for those countries.

Figure 7. Myers index for 10 Arab countries, most recent census



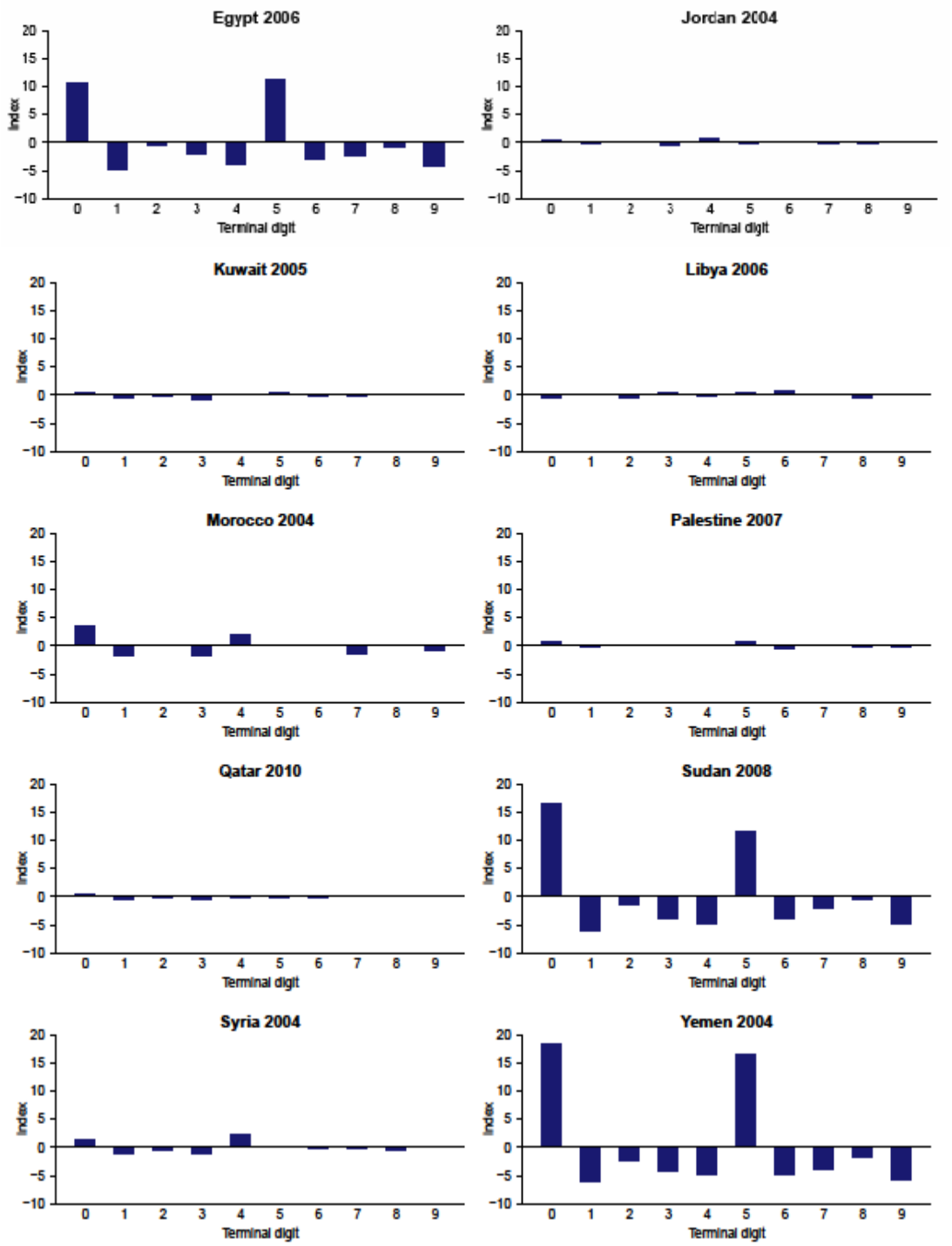
Note: See annex table AI.8 for disaggregation by sex.

Bachi index

Like the other indices, the Bachi index suggests that the age data from Egypt, the Sudan and Yemen are of poor quality, and that age misreporting is minimal in Jordan, Kuwait, Morocco, Palestine, Qatar and the Syrian Arab Republic (figure 8). In all countries, there is at least some repulsion from ages ending in 1, 2, 3, 6, 7, 8 and 9. Accordingly, the attraction to ages ending with 0 or 5 is apparent (though slight) in

Jordan, Kuwait, Palestine, Qatar and the Syrian Arab Republic, and severe in Egypt, the Sudan and Yemen. When disaggregated by sex, both the Myers and Bachi indices reveal less distortion in the reporting of men's ages than women's for all countries (see annex table AI.9 for Bachi index results disaggregated by sex).

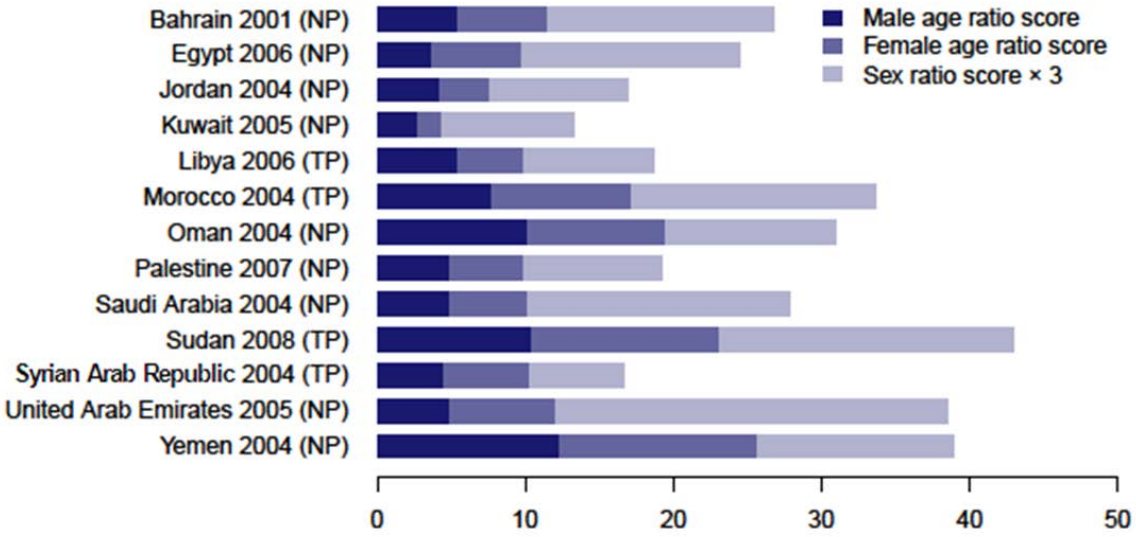
Figure 8. Bachi index for 10 Arab countries, most recent census



United Nations age-sex accuracy index (UNACI).

The UNACI score is less than 20 for Jordan, Kuwait, Palestine and the Syrian Arab Republic, suggesting that grouped age data are of good quality (figure 9). Grouped data from Bahrain, Egypt, Morocco, Oman, Saudi Arabia, the United Arab Emirates and Yemen are of relatively good quality. However, data from the Sudan appear to be of poor quality, with a UNACI score over 40.

Figure 9. United Nations age-sex accuracy index for 13 Arab countries, most recent census

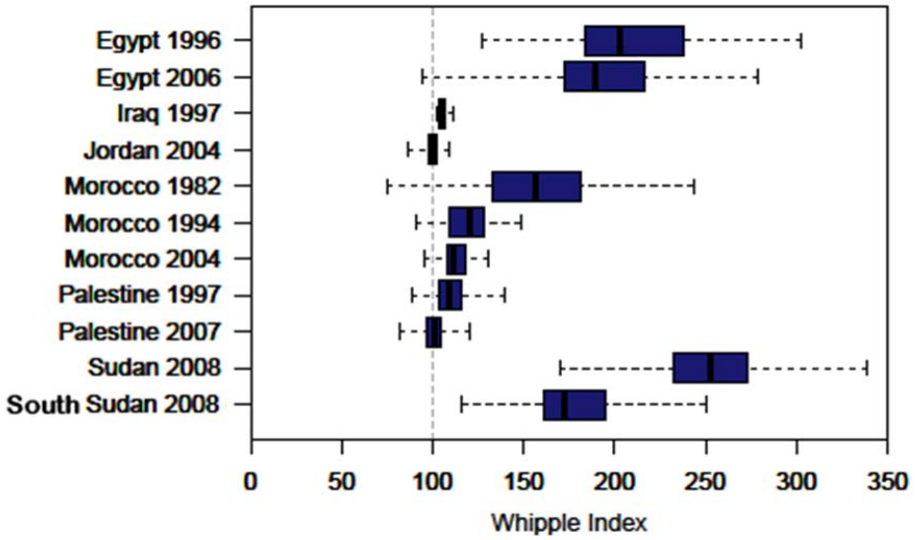


Note: (NP) indicates national population; (TP) indicates total population. See annex table AI.10 for the exact UNACI values.

IV. ANALYSIS OF AGE HEAPING AT THE SUBNATIONAL LEVEL IN RECENT POPULATION CENSUSES

The spread of district-level Whipple indices varies considerably across censuses.[15] Whipple indices from censuses in which severe age heaping exists tend to have much wider spreads than countries whose overall Whipple index is closer to 100. Censuses can be put into two groups: those with high, variable Whipple indices and with Whipple indices that stay within a narrow range close to 100 (figure 10). The censuses of Egypt (1996 and 2006), Morocco (1982), the Sudan (2008) and South Sudan (2008) belong to the former group. The censuses of Iraq (1997), Jordan (2004), Morocco (1994 and 2004) and Palestine (1997 and 2007) belong to the latter group. A dramatic decrease in age heaping is seen between the 1982 and 1994 censuses of Morocco.

Figure 10. Boxplots of district-level Whipple indices



Summary statistics for predictors also vary considerably from country to country (table 4). According to the 2004 census of Jordan, 90.5 per cent of the population was literate. The 2007 census of Palestine reported 93.5 per cent literacy, while the 2008 census of South Sudan reported only 26 per cent literacy. In general, higher literacy rates also have smaller standard deviations.

Per cent urban varies considerably among districts, as one might expect. Since the urban/rural determination is made at the district level in some of the census files, many districts are either 0 per cent urban or 100 per cent urban by construction.

Per cent female varies very little, and ranges from 48.8 per cent in Jordan (2004) to 51.0 per cent in South Sudan (2008). Per cent female is considerably more variable in the 2008 censuses of both the Sudan and South Sudan than it is in any of the other census files (both have a standard deviation of 3 per cent or higher, while the next largest standard deviation is 1.2 per cent).

The 2008 census of South Sudan records the lowest median age at 15 years old, while Morocco (2004) has the highest at 24 years old. Standard deviations of median age are fairly consistent across censuses (generally close to two years).

[15] Subnational census data for Egypt, Iraq, Jordan, Morocco, Palestine and the Sudan are provided in annex II.

TABLE 4. SUMMARY STATISTICS FOR PREDICTORS

Country	Per cent literate	Per cent urban	Per cent female	Median age	Per cent in camps
Egypt 1996	55.6 (16.0)	42.6 (41.4)	48.9 (0.9)	20 (2.8)	-
Egypt 2006	66.2 (12.8)	42.4 (41.2)	48.9 (0.9)	22 (2.6)	-
Iraq 1997	73.8 (9.9)	67.4 (24.8)	50.1 (0.5)	17 (1.7)	-
Jordan 2004	90.5 (3.9)	78.7 (27.5)	48.8 (0.6)	20 (1.8)	-
Morocco 1982	30.4 (14.3)	..	50.4 (1.1)	18 (1.5)	-
Morocco 1994	42.1 (15.0)	..	50.4 (1.2)	20 (2.3)	-
Morocco 2004	52.4 (13.4)	..	50.7 (1.2)	24 (2.4)	-
Palestine 1997	86.3 (3.4)	53.1 (46.3)	49.2 (0.9)	16 (2.0)	16.1 (32.6)
Palestine 2007	93.5 (2.2)	71.9 (36.3)	49.2 (0.7)	17 (2.4)	9.9 (26.3)
Sudan 2008	56.7 (20.3)	33.2 (33.1)	50.4 (3.0)	18 (2.7)	-
South Sudan 2008	26.0 (18.8)	24.4 (25.3)	51.0 (3.3)	15 (2.0)	-

Note: A hyphen (-) indicates that the item is not applicable. Two dots (..) indicate that data are not available. Standard deviations (weighted by district counts) are shown in parentheses.

Where severe age heaping is present, it appears to be highly correlated with predictor variables, with the exception of South Sudan (2008) (table 5). A look at pairwise correlations between the Whipple index and each of the predictors reveals that literacy is most strongly correlated with age heaping in most censuses. Wealth also appears to be strongly correlated with heaping. The correlation coefficients between literacy and the Whipple index are curiously positive, though small, in Jordan (2004), Palestine (2007) and South Sudan (2008), meaning there are weak associations between increased literacy and poorer quality age data at the district level in these censuses.

TABLE 5. PAIRWISE CORRELATIONS WEIGHTED BY DISTRICT COUNTS BETWEEN THE WHIPPLE INDEX AND PREDICTORS

Country	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Per cent in camps
Egypt 1996	-0.84	-0.73	-0.67	-0.02	-0.69	-
Egypt 2006	-0.84	-0.73	-0.67	-0.04	-0.69	-
Iraq 1997	-0.47	-0.43	-0.41	0.28	-0.45	-
Jordan 2004	0.18	0.14	0.01	0.15	0.14	-
Morocco 1982	-0.84	-0.80	..	0.07	-0.68	-
Morocco 1994	-0.77	-0.77	..	-0.16	-0.68	-
Morocco 2004	-0.58	-0.58	..	-0.18	-0.57	-
Palestine 1997	-0.15	-0.03	-0.06	-0.09	-0.43	0.32
Palestine 2007	0.12	-0.03	-0.11	-0.06	0.09	0.23
Sudan 2008	-0.78	-0.79	-0.65	0.30	-0.80	-
South Sudan 2008	0.30	0.35	0.26	0.06	-0.04	-

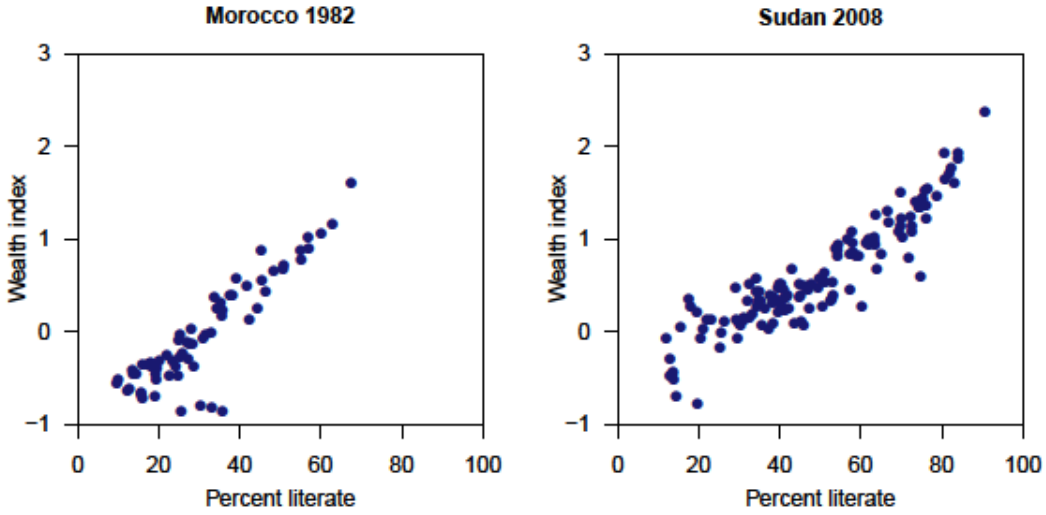
Note: A hyphen (-) indicates that the item is not applicable. Two dots (..) indicate that data are not available.

The sign of the correlation coefficients between the Whipple index and median age is also unexpected for most censuses. Although older individuals are more likely to misreport their ages than younger individuals, at the district level, older-age areas are more likely to have better quality age data than younger-age areas.

Interestingly, correlation coefficients between the Whipple index and five of the six explanatory variables are the same to two decimal places for data from the censuses of Egypt of 1996 and 2006.

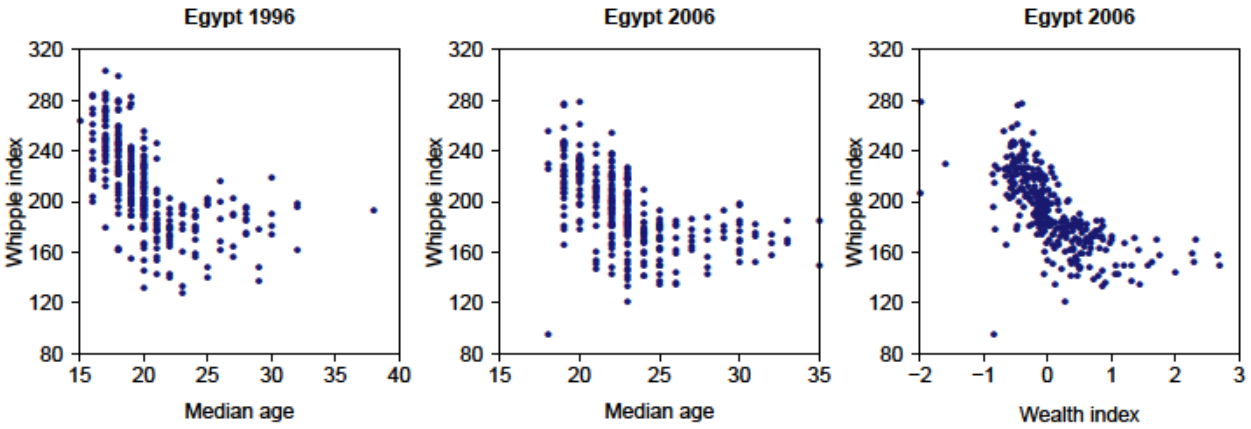
Many of the predictors considered for the models are correlated with each other, which causes issues with collinearity for a few of the models. In particular, literacy and wealth index tend to be strongly, positively correlated with each other. In the models of the 1982 census of Morocco and the 2008 census of the Sudan, wealth is removed, because it complicates the interpretation of coefficients and inflates their standard errors. In both of these censuses, there is a greater correlation between wealth and literacy than there is between either wealth or literacy and the Whipple index. Figure 11 illustrates the strong correlation between wealth and literacy in the data. Thus, either wealth or literacy can be used. This study has chosen literacy simply because it does a better job of explaining the variation in the Whipple index.

Figure 11. Scatterplots of per cent literate vs. wealth index in the 1982 census of Morocco and the 2008 census of the Sudan



In most of the models, the predictors that are significant appear to have linear relationships with the response variable. Only a small number of potentially nonlinear relationships surfaced, all in the Egypt censuses (figure 12). In the 1996 census of Egypt, median age appears to have a nonlinear relationship with the Whipple index. In the 2006 census of Egypt, both median age and wealth index appear to have nonlinear relationships with the Whipple index. Adding squared terms in the models for each of these predictors does increase the significance of the variables, but does not greatly improve the overall fit of the model. The squared terms are therefore left out of the final models.

Figure 12. Scatterplots of nonlinear relationships in the 1996 and 2006 censuses of Egypt



Note: The outliner in the Egypt 1996 plot indicating a median age of 38, represents a group of districts in the Cairo governorate with populations under 20,000.

Models for the censuses of Egypt (1996 and 2006), Morocco (1982) and the Sudan (2008) explain variation in Whipple index quite well; adjusted R^2 values for each of these models is greater than 0.7 (table 6). Morocco (1994) also deserves mention, with an adjusted R^2 of 0.61. The models for Iraq (1997), Jordan (2004), Palestine (2007) and South Sudan (2008) all fail to adequately explain changes in Whipple index. It is not surprising that Whipple indices that are close to 100 are difficult to model, but the 2008 census of South Sudan stands alone as the only census whose age distribution is characterized by severe age heaping that does not lend itself to being properly modelled here.

TABLE 6. EFFECTS OF DISTRICT CHARACTERISTICS ON THE WHIPPLE INDEX

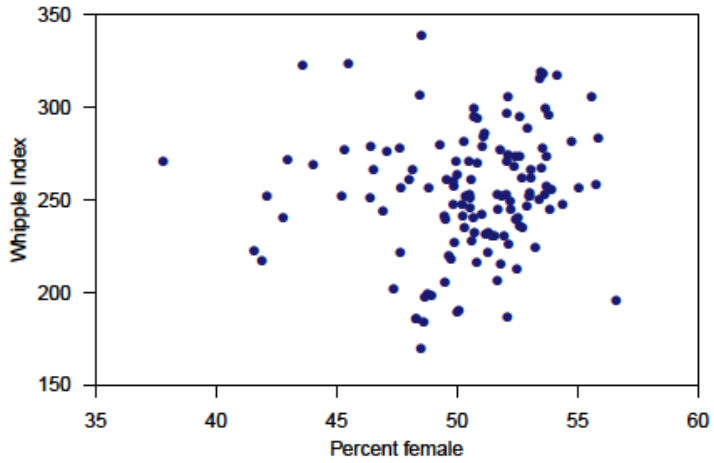
	Egypt 1996	Egypt 2006	Morocco 1982	Morocco 1994	Morocco 2004	Palestine 1997	Sudan 2008
Intercept	442.68* (66.73)	226.62* (46.49)	234.31 (104.20)	215.05* (47.64)	171.58* (31.67)	304.09* (74.36)	195.85* (39.94)
Per cent literate	-2.30* (0.20)	-1.67* (0.14)	-1.76* (0.24)	-0.38 (0.21)	-0.07 (0.18)	-1.10 (0.53)	-1.04* (0.15)
Wealth index	17.53* (4.55)	1.76 (3.53)	..	-8.37 (6.71)	-3.88 (5.12)	8.07 (4.25)	..
Per cent urban	-0.07 (0.04)	-0.09 (0.03)	..	..	..	0.02 (0.03)	0.00 (0.07)
Per cent female	-1.93 (1.35)	1.80 (0.94)	0.15 (1.92)	-1.62 (0.84)	-0.82 (0.61)	-1.71 (1.08)	3.18* (0.65)
Median age	-0.22 (0.76)	-0.11 (0.61)	-1.35 (2.25)	0.13 (0.92)	-0.57 (0.56)	-1.15 (0.57)	-3.16 (1.05)
Per cent in camps	-	-	-	-	-	0.09 (0.04)	-
Adjusted R^2	0.72	0.72	0.70	0.61	0.34	0.25	0.75

Note: A hyphen (-) indicates that the item is not applicable. Two dots (..) indicate that data are not available. An asterisks (*) denotes effects that are significant at the 0.05 level after *p*-values have been adjusted to correct for multiple testing.

Per cent literate is significant in all of the best-performing models. Wealth index is only significant for Egypt (1996), despite being strongly correlated with the Whipple index in many of the censuses. Even in the Egypt (1996) model, the coefficient for wealth is positive, which is counterintuitive, and likely a result of collinearity.

Per cent female is only significant in the model of the Sudan 2008. It is probably no coincidence that per cent female varies more in the census of the Sudan than it does in the census of any of the other countries. Still, a scatterplot of per cent female and the Whipple index does not appear to show a strong linear relationship (figure 13). Nevertheless, per cent female is identified as a significant predictor in the model.

Figure 13. Scatterplot of per cent female vs. Whipple index in the 2008 census of the Sudan



In general, the model performs well for censuses with more severe age heaping, and poorly for those with only minor heaping. Literacy and wealth tend to track each other closely in most censuses, which complicates interpretation of effects when both are included in the model. Per cent of a district that is female is consistently either weakly correlated or uncorrelated with the severity of age heaping in the district. Model coefficients for literacy are negative, meaning as literacy increases, the Whipple index score for the district decreases.

V. CONCLUSIONS AND RECOMMENDATIONS

This report has shown that although census-taking dates back to the mid-nineteenth century in the Arab region, age misreporting continues to be a problem, which affects derivations of other population characteristics. Age and sex data collected from the census should therefore be evaluated for possible errors, not only age reporting errors, but also possible under-enumeration of certain population groups.

The indices presented in this report for the analysis of age-sex data can be used to identify specific types of errors that are likely to impact the reliability of collected data for specific population groups. Although the analysis of age-sex distribution can reveal errors, such analysis cannot differentiate between content and coverage errors. Data from other sources are needed to verify estimates of particular types of errors.[16]

The evaluation of the quality and reliability of age and sex data from Arab countries revealed some issues in the reporting of ages. Those issues include the preference for ages that are multiples of 5 and an avoidance of other digits. Age reporting errors are more severe for women than they are for men in most Arab countries. Problems related to age reporting can be mitigated by well-kept civil registers, but these are not available in many countries. In the short term, it is necessary to improve the methods of collecting age data and to assess the impact of errors on the measurement of demographic phenomena.

Findings from the analysis of district-level age data suggest that the use of a linear model can be justified when severe age heaping is present in a census. Per cent literate is the best predictor of age heaping, followed closely by wealth. Per cent female does not appear to explain any additional variation in the response variable when wealth and literacy are also in the model, perhaps because per cent female does not vary significantly from district to district. Model improvements could be made by considering other background variables, or using 100 as a floor for predictions, since values of the Whipple index less than 100 no longer represent an improvement to data quality. Differences between countries in census dates, number of districts and methods of data collection may limit the usefulness of country comparisons.

Further work is needed to evaluate the quality of demographic data in Arab censuses. Little is known, for example, about the quality of birth and death reporting, as well as the quality of data collected on other demographic variables such as marriage and divorce, household and family composition, and internal and international migration. Work should also be extended to include other Arab countries for which census microdata are not yet available. Finally, the quality of census data in the Arab region should also be compared to other developing regions of the world, including Central Asia, Latin America and sub-Saharan Africa.

[16] United Nations, 2011.

Annex I

NATIONAL DATA

TABLE AI.1. MODERN POPULATION CENSUSES COMPLETED AND PLANNED IN ARAB COUNTRIES (1930-PRESENT)

Country (year of independence)	Period							
	Before 1950	1950-1959	1960-1969	1970-1979	1980-1989	1990-1999	2000-2009	2010-present
Algeria (1962)	1948	1954	1967	1977	1987	1998	2008	(2016)
Bahrain (1971)	--	1950 1959	1965	1971	1981	1991[a/]	2001[a/]	2010[a/]
Egypt (1922)	1947	--	1960	1976	1986	1996	2006	--
Iraq (1932)	1947[b/]	1957	1965/66[c/]	1977	1987	1997	--	--
Jordan (1946)	--	1952	1961	1979	--	1994	2004	(2015)
Kuwait (1961)	--	1957	1961 1965	1970 1975	1980 1985	1995	2005	2011
Lebanon (1943)	1932	--	--	--	--	--	--	--
Libya (1951)	1936	1954	1964	1973	1984	1995	2006	--
Mauritania (1960)	1944	--	--	1977[d/]	1988	--	2000	2013
Morocco (1956)	--	--	1960	1971	1982	1994	2004	(2014)
Oman (1970)	--	--	--	--	--	1993	2003	2010
Palestine			1967[e/]		--	1997	2007	--
Qatar (1971)	--	--	--	1970	1986	1997	2004	2010
Saudi Arabia (1932)	--	--	1962/63[c/]	1974[f/]	--	1992	2004	2010
Sudan (1956)	--		--	1973	1983	1993[g/]	2008	--
Syrian Arab Republic (1946)	1947[b/]	--	1960	1970 1974	1980	1994	2004	..
Tunisia (1956)	1946	1956	1966	1975	1984	1994	2004	(2014)
United Arab Emirates (1971)	--	--	1968	1970	1980 1985	1995	2005	--
Yemen (1968)*	--	--	--	1973 1975	1986 1988	1994	2004	(2014)

Sources: National statistical offices, the United States Department of Commerce, the United States Census Bureau (http:/www.census.gov) and the United Nations Statistics Division (http://unstats.un.org/unsd/demographic/sources/census/censusdates.htm).

a/ Population figures compiled from administrative registers.
b/ Quick count.
c/ Incomplete count.
d/ General census for sedentary population and sample census for nomads.
e/ Supervised by the Israeli Central Bureau of Statistics (ICBS) in the West Bank and Gaza Strip under Israeli control.
f/ Not reliable.
g/ North only.

* Enumerations of the former Democratic Yemen (South) were conducted in 1973 and 1988. Enumerations of the former Yemen Arab Republic (North) were conducted in 1975 and 1986. They became one country in May 1990.

Note: Two hyphens (--) indicate no census was taken. Two dots (..) denote an expected census during the decade; the year of a planned census is given between parentheses.

TABLE AI.2. SEX RATIO BY FIVE-YEAR AGE GROUPS FOR THE MOST RECENT CENSUS

Country	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65+
Bahrain 2010*	104	105	105	105	106	105	101	96	93	93	94	118	117	93
Egypt 2006*	105	106	108	105	105	98	105	99	102	105	102	116	112	109
Iraq 1997	102	104	104	104	99	97	96	85	95	94	103	103	86	79
Jordan 2004*	105	105	105	106	107	111	108	108	106	105	102	96	112	102
Kuwait 2005*	105	103	106	104	103	100	96	93	92	92	84	81	74	88
Libya 2006	95	96	96	97	98	97	98	100	100	98	105	99	86	94
Morocco 2004	104	103	103	99	94	92	92	90	92	104	104	96	85	94
Oman 2010*	104	103	105	105	105	101	99	103	100	97	89	91	101	105
Palestine 2007*	105	105	105	104	105	103	102	104	108	107	100	103	86	72
Qatar 2010	105	105	105	127	405	368	401	483	506	450	386	364	261	144
Saudi Arabia 2004*	101	101	94	101	97	103	99	99	110	113	111	95	94	106
Sudan 2008	106	109	115	106	97	89	93	96	104	112	113	123	123	130
Syrian Arab Republic 2004	106	106	107	105	104	103	102	103	105	107	108	110	103	110
United Arab Emirates 2005*	105	105	107	106	96	96	98	94	89	90	103	125	136	142
Yemen 2004*	105	105	112	104	104	98	102	90	95	94	104	109	111	112

Source: ESCWA calculations based on national census data.

* National population.

TABLE AI.3. HISTORY OF OFFICIAL CENSUSES

Country	Date of first census	Date of most recent census	Total number of censuses
Egypt	3 May 1882	21 November 2006	13
Iraq	1927	16 October 1997	7
Jordan	November 1961	2 October 2004	5
Kuwait	20 April 1965	20 April 2011	8
Morocco	1960	1 November 2004	5
Oman	1 December 1993	12 December 2010	3
Palestine	10 December 1997	1 December 2007	2
Qatar	March 1986	20 April 2010	4
Syrian Arab Republic	September 1960	14 September 2004	5
United Arab Emirates	31 December 1975	6 December 2005	5
Yemen	1994	17 December 2004	4

Source: Responses from member countries to ESCWA questionnaire.

TABLE AI.4. METHODOLOGY, MOST RECENT CENSUS

Country	Duration of census (days)	Traditional	PAPI	CAPI	Number of questionnaires	GPS	Stage of GPS/GIS use		
							Preparation	Enumeration	Reporting
Egypt	53	Yes	Yes	No	1	Yes	Yes	Yes	Yes
Iraq	1	Yes	Yes	No	1	Yes	Yes	No	Yes
Jordan	10	Yes	Yes	No	4	No	--	--	--
Kuwait	40	Yes	Yes	No	1	Yes	Yes	Yes	No
Morocco	20	Yes	Yes	No	3	No	--	--	--
Oman	10	No	No	Yes (PDA)	1	Yes	Yes	Yes	Yes
Palestine	15	Yes	Yes	No	1	No	--	--	--
Qatar	15	Yes	No	No	7	Yes	Yes	Yes	Yes
Syrian Arab Republic	7	Yes	Yes	No	1	Yes	Yes	No	No
United Arab Emirates	10	Yes	No	Yes (PDA)	1	Yes	Yes	No	Yes
Yemen	10	Yes	Yes	No	2	Yes	Yes	Yes	Yes

Source: Responses from member countries to ESCWA questionnaire.

TABLE AI.5. EVALUATION METHODOLOGY

Country	Post-enumeration survey	Demographic analysis	Comparison with administration register	Comparison with surveys
Egypt	Yes	Yes	No	Yes
Iraq	Yes	Yes	No	Yes
Jordan	Yes	No	No	No
Kuwait	No	No	Yes	No
Morocco	No	No	No	No
Oman	No	Yes	Yes	Yes
Palestine	Yes	Yes	Yes	No
Qatar	No	Yes	Yes	No
Syrian Arab Republic	No	Yes	No	Yes
United Arab Emirates	No	No	Yes	No
Yemen	No	Yes	No	Yes

Source: Responses from member countries to ESCWA questionnaire.

TABLE AI.6. UNDERCOVERAGE AND REFUSAL RATES, MOST RECENT CENSUS
(*Percentage*)

Country	Undercoverage rate	Refusal rate
Egypt	8.3	..
Iraq	..	..
Jordan	4.0	..
Kuwait	5.0	0
Morocco	..	..
Oman	..	..
Palestine	2.7	..
Qatar	5.0	..
Syrian Arab Republic	..	..
United Arab Emirates	5.0	0
Yemen	..	..

Source: Responses from member countries to ESCWA questionnaire.

Note: Two dots (..) indicate that the data are not available.

TABLE AI.7. WHIPPLE INDEX BY SEX

Country	Men	Women	Total
Bahrain 2001	100	101	100
Egypt 1996	184	240	212
Egypt 2006	175	218	197
Iraq 1997	104	105	105
Jordan 2004*	100	101	101
Kuwait 2005*	105	105	105
Libya 2006	100	101	100
Morocco 1982	147	180	164
Morocco 1994	110	129	120
Morocco 2004	108	118	113
Palestine 1997*	103	115	109
Palestine 2007*	100	102	101
Qatar 2010	102	102	102
Sudan 2008	218	235	227
Syrian Arab Republic 2004	105	107	106
Yemen 2004*	249	271	260

Source: ESCWA calculations based on national census data.

* National population.

TABLE AI.8. MYERS INDEX BY SEX, MOST RECENT CENSUS

Country		Terminal digit										Index
		0	1	2	3	4	5	6	7	8	9	
	Men	6.1	-2.8	0.2	-1.2	-2.9	6.3	-1.7	-1.5	-0.5	-1.9	25.2
	Women	9.8	-3.9	-0.6	-2.4	-3.8	9.4	-3.0	-2.4	-0.7	-2.4	38.5
Egypt 2006*	Total	7.9	-3.4	-0.2	-1.8	-3.4	7.8	-2.3	-1.9	-0.6	-2.1	31.5
	Men	0.2	-0.1	0.3	-0.3	0.5	-0.2	0.0	-0.2	-0.1	0.0	1.9
	Women	0.4	-0.2	0.2	-0.5	0.7	-0.1	0.1	-0.4	-0.2	0.1	2.9
Jordan 2004*	Total	0.3	-0.2	0.2	-0.4	0.6	-0.1	0.0	-0.3	-0.1	0.0	2.3
	Men	-0.2	-0.7	-0.4	-0.2	0.5	1.0	0.3	0.1	-0.1	-0.3	3.7
	Women	-0.1	-0.5	-0.1	-0.3	0.6	0.8	0.1	-0.1	-0.3	0.0	3.1
Kuwait 2005*	Total	-0.2	-0.6	-0.3	-0.3	0.6	0.9	0.2	0.0	-0.2	-0.2	3.4
	Men	-0.2	0.1	-0.5	0.1	-0.4	0.2	0.6	0.2	-0.2	0.1	2.7
	Women	-0.4	-0.1	-0.5	0.2	-0.1	0.3	0.7	0.1	-0.2	0.0	2.5
Libya 2006	Total	-0.3	0.0	-0.5	0.2	-0.2	0.3	0.7	0.1	-0.2	0.0	2.5
	Men	1.9	-1.1	0.4	-1.2	1.5	-0.1	0.0	-0.8	0.2	-0.7	7.9
	Women	4.0	-1.8	0.0	-1.9	1.6	0.2	0.1	-1.4	0.2	-1.0	12.2
Morocco 2004	Total	3.0	-1.5	0.2	-1.6	1.5	0.0	0.1	-1.1	0.2	-0.8	10.0
	Men	-0.1	-0.2	0.0	0.1	0.2	0.4	-0.1	0.2	-0.2	-0.3	2.0
	Women	0.7	-0.6	-0.1	-0.1	0.0	1.0	-0.5	0.2	-0.2	-0.4	3.7
Palestine 2007*	Total	0.3	-0.4	-0.1	0.0	0.1	0.7	-0.3	0.2	-0.2	-0.4	2.6
	Men	0.1	-0.2	0.3	0.3	0.5	0.2	-0.1	-0.3	-0.5	-0.4	3.0
	Women	-0.1	-0.6	-0.2	0.0	0.4	0.4	0.4	0.2	-0.4	-0.2	3.0
Qatar 2010	Total	0.1	-0.3	0.2	0.3	0.5	0.3	0.0	-0.2	-0.4	-0.4	2.6
	Men	11.7	-5.0	0.6	-2.9	-3.5	7.7	-2.7	-1.8	0.5	-4.5	41.0
	Women	13.0	-5.6	0.3	-3.7	-4.0	9.1	-3.1	-1.8	0.3	-4.5	45.4
Sudan 2008	Total	12.4	-5.3	0.5	-3.3	-3.7	8.4	-2.9	-1.8	0.4	-4.5	43.3
	Men	0.6	-0.7	-0.1	-0.9	1.5	0.0	-0.2	-0.1	-0.3	0.2	4.7
Syrian Arab	Women	1.2	-1.0	-0.2	-1.4	1.9	-0.1	-0.3	-0.3	-0.4	0.6	7.5
Republic 2004	Total	0.9	-0.9	-0.1	-1.2	1.7	0.0	-0.3	-0.2	-0.3	0.4	6.0
	Men	11.7	-4.7	-0.3	-3.0	-3.2	10.1	-3.3	-2.7	-0.1	-4.4	43.5
	Women	13.6	-5.3	-1.5	-3.7	-3.7	11.8	-3.8	-3.1	0.2	-4.5	51.3
Yemen 2004*	Total	12.6	-5.0	-0.9	-3.4	-3.4	10.9	-3.5	-2.9	0.0	-4.5	47.2

Source: ESCWA calculations based on national census data.

* National population.

TABLE AI.9. BACHI INDEX BY SEX, MOST RECENT CENSUS

Country		Terminal digit										Index
		0	1	2	3	4	5	6	7	8	9	
	Men	8.1	-4.0	0.1	-1.4	-3.2	9.3	-2.2	-1.8	-0.7	-3.8	17.3
	Women	13.3	-5.4	-1.2	-2.9	-4.4	13.3	-3.8	-3.1	-1.1	-4.4	26.5
Egypt 2006*	Total	10.6	-4.7	-0.5	-2.2	-3.8	11.3	-3.0	-2.4	-0.9	-4.1	21.7
	Men	0.2	-0.2	0.2	-0.4	0.6	-0.1	0.1	-0.1	-0.1	0.1	1.0
	Women	0.3	-0.4	0.0	-0.5	0.9	0.0	0.3	-0.3	-0.3	0.2	1.6
Jordan 2004*	Total	0.3	-0.3	0.1	-0.5	0.7	-0.1	0.2	-0.2	-0.2	0.1	1.3
	Men	0.3	-0.4	-0.2	-0.7	-0.1	0.5	-0.1	-0.1	0.1	0.2	1.3
	Women	0.3	-0.4	-0.1	-0.6	0.3	0.6	-0.1	-0.3	-0.2	0.3	1.6
Kuwait 2005*	Total	0.3	-0.4	-0.1	-0.7	0.1	0.5	-0.1	-0.2	0.0	0.2	1.3
	Men	-0.2	0.3	-0.6	0.2	-0.4	0.3	0.7	0.2	-0.4	0.0	1.7
	Women	-0.5	0.1	-0.6	0.3	-0.1	0.4	0.8	0.0	-0.3	0.0	1.5
Libya 2006	Total	-0.4	0.2	-0.6	0.3	-0.2	0.3	0.7	0.1	-0.4	0.0	1.6
	Men	2.4	-1.3	0.4	-1.3	1.7	-0.1	0.1	-1.1	0.1	-0.6	4.5
	Women	4.7	-2.1	-0.1	-2.1	1.9	0.4	0.1	-1.8	0.0	-1.0	7.1
Morocco 2004	Total	3.6	-1.7	0.1	-1.7	1.8	0.2	0.1	-1.5	0.0	-0.8	5.8
	Men	0.1	-0.1	0.1	0.3	0.3	0.4	-0.3	0.2	-0.3	-0.3	1.2
	Women	1.1	-0.5	0.0	0.1	0.1	1.0	-0.9	0.0	-0.3	-0.3	2.2
Palestine 2007*	Total	0.6	-0.3	0.0	0.2	0.2	0.7	-0.6	0.1	-0.3	-0.3	1.7
	Men	0.5	-0.4	-0.1	-0.4	0.0	-0.1	-0.2	-0.1	0.2	0.2	1.1
	Women	0.5	-0.2	0.0	-0.6	-0.2	-0.2	-0.1	0.1	0.0	0.2	1.1
Qatar 2010	Total	0.5	-0.4	-0.1	-0.4	-0.1	-0.1	-0.2	0.0	0.2	0.2	1.1
	Men	15.9	-5.8	-1.1	-3.6	-4.6	10.4	-3.8	-2.0	-0.5	-4.9	26.3
	Women	17.3	-6.5	-1.8	-4.4	-5.1	12.2	-4.1	-2.1	-0.6	-4.9	29.6
Sudan 2008	Total	16.6	-6.2	-1.5	-4.0	-4.9	11.4	-4.0	-2.1	-0.6	-4.9	28.0
	Men	1.0	-0.9	-0.3	-0.9	2.0	0.3	-0.1	-0.1	-0.5	-0.1	3.2
Syrian Arab Republic	Women	1.5	-1.2	-0.5	-1.5	2.6	0.2	-0.3	-0.3	-0.7	0.4	4.5
2004	Total	1.2	-1.1	-0.4	-1.2	2.3	0.2	-0.2	-0.2	-0.6	0.1	3.8
	Men	17.4	-5.9	-1.7	-4.0	-4.5	15.4	-4.4	-3.7	-1.9	-5.8	32.3
	Women	19.1	-6.5	-3.2	-4.6	-5.1	17.7	-5.3	-4.3	-1.5	-5.6	36.5
Yemen 2004*	Total	18.3	-6.2	-2.5	-4.3	-4.8	16.5	-4.9	-4.0	-1.7	-5.7	34.4

Source: ESCWA calculations based on national census data.

* National population.

TABLE AI.10. UNITED NATIONS AGE-SEX ACCURACY INDEX BY COMPONENT

Country	Index	Component		
		Age ratio score males	Age ratio score females	Sex ratio score × 3
		(*Percentage*)		
Bahrain 2001*	26.8	20	23	57
Egypt 2006*	24.5	15	25	60
Jordan 2004*	16.9	25	20	55
Kuwait 2005*	13.3	20	13	67
Libya 2006	18.7	29	24	47
Morocco 2004	33.6	23	28	49
Oman 2004*	30.9	33	30	37
Palestine 2007*	19.2	25	26	49
Saudi Arabia 2004*	27.9	18	19	63
Sudan 2008	42.9	24	29	47
Syrian Arab Republic 2004	16.6	27	35	38
United Arab Emirates 2005*	38.5	13	19	68
Yemen 2004*	38.9	32	34	34

Source: ESCWA calculations based on national census data.

* National population.

Annex II

SUBNATIONAL DATA

TABLE AII.1. GOVERNORATE-LEVEL DESCRIPTIVE STATISTICS FOR THE 1996 CENSUS OF EGYPT

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Population
Alexandria	183	72	0.65	100	49	24	3 339 062
Asiut	247	41	-0.61	27	49	17	2 800 930
Aswan	211	60	0.04	43	50	19	973 756
Bani Swif	231	41	-0.83	24	49	17	1 859 207
Behera	215	46	-0.21	23	49	19	3 993 016
Cairo	179	73	0.76	100	49	24	6 800 958
Dakahlia	204	58	0.24	28	49	20	4 223 525
Demietta	217	62	0.41	27	49	21	913 551
Fayoum	247	37	-1.03	22	48	17	1 989 765
Gharbia	204	60	0.04	31	50	20	3 405 720
Giza	208	61	0.33	54	48	20	4 784 077
Ismailia	203	64	0.31	50	49	20	714 825
Kafr Sheikh	231	46	-0.16	23	50	19	2 223 030
Kaliobia	199	60	0.07	41	49	20	3 301 227
Luxor	263	51	0.02	46	49	20	361 136
Marsa Matroh	208	48	-0.38	56	48	18	211 738
Menia	220	39	-1.05	19	49	18	3 310 114
Menoufia	210	58	-0.33	20	49	20	2 760 419
New Valley	148	70	0.36	48	49	20	141 773
North Sinai	200	59	-0.33	59	48	18	253 556
Port Said	181	76	0.74	100	49	24	472 333
Qena	261	41	-0.47	21	50	18	2 440 695
Red Sea	187	70	0.41	75	46	20	157 314
Sharkia	232	52	0.00	23	49	19	4 280 727
Sohag	250	40	-0.65	22	49	18	3 122 799
South Sinai	194	64	0.28	54	42	20	54 826
Suez	157	72	0.74	100	49	21	417 526

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.2. GOVERNORATE-LEVEL DESCRIPTIVE STATISTICS FOR THE 2006 CENSUS OF EGYPT

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Population
Alexandria	184	78	0.36	99	49	25	4 107 330
Asiut	222	54	-0.41	26	49	19	3 430 020
Aswan	200	73	-0.08	42	50	22	1 178 320
Bani Swif	215	53	-0.56	23	49	20	2 287 820
Behera	211	58	-0.24	19	49	22	4 745 440
Cairo	165	79	0.73	100	49	25	7 821 590
Dakahlia	191	68	0.13	28	49	23	4 976 190
Damietta	187	75	0.11	38	49	23	1 092 240
Fayoum	220	53	-0.43	22	48	20	2 509 870
Gharbia	197	71	0.02	30	49	23	4 004 370
Giza	191	69	0.22	59	48	22	6 259 080
Ismailia	172	74	0.13	45	49	22	949 310
Kafr Sheikh	221	61	-0.13	23	49	22	2 615 750
Kaliobia	179	69	0.03	38	49	23	4 233 080
Luxor	212	67	0.00	47	50	22	446 170
Marsa Matroh	212	59	-0.43	70	48	19	319 670
Menia	216	52	-0.55	19	49	20	4 149 680
Menoufia	194	69	-0.15	20	49	22	3 261 950

TABLE AII.2 (*continued*)

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Population
New Valley	144	79	0.11	48	49	23	182 780
North Sinai	171	72	-0.50	60	48	20	342 740
Port Said	155	82	0.37	100	49	25	558 610
Qena	224	59	-0.42	21	50	20	2 986 810
Red Sea	176	82	0.03	95	45	24	233 640
Sharkia	198	63	-0.07	23	49	22	5 346 820
Sohag	238	55	-0.38	21	49	20	3 740 100
South Sinai	173	77	0.10	52	42	24	87 010
Suez	147	80	0.48	100	49	23	508 990

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.3. GOVERNORATE-LEVEL DESCRIPTIVE STATISTICS FOR THE 1997 CENSUS OF IRAQ

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Population
Al-Anbar	107	71	-0.01	53	50	15	1 026 960
Al-Basrah	106	74	0.06	81	50	17	1 482 460
Al-Muthanna	108	55	-0.58	47	51	16	446 120
Al-Najaf	107	71	0.05	70	50	17	784 530
Al-Qadisiya	105	64	-0.43	53	51	16	756 380
Al-Tameem	105	75	0.12	71	50	17	762 650
Babylon	105	76	-0.27	48	50	17	1 188 740
Baghdad	103	83	0.27	89	50	19	5 475 400
Diala	105	75	-0.04	41	50	17	1 206 800
Kerbela	107	75	-0.02	67	50	17	596 950
Maysan	105	57	-0.45	66	51	16	644 060
Nineveh	104	72	-0.11	62	50	15	2 060 260
Salah Al-Deen	105	69	-0.31	45	51	15	915 700
Thi-Qar	106	62	-0.45	59	51	16	1 200 980
Wasit	104	64	-0.44	53	51	17	792 260

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.4. GOVERNORATE-LEVEL DESCRIPTIVE STATISTICS FOR THE 2004 CENSUS OF JORDAN

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Population
Ajloun	103	88	-0.20	66	50	18	118 690
Amman	101	93	0.33	92	49	21	1 900 540
Aqaba	99	88	-0.06	87	47	18	93 130
Balqa	104	87	-0.08	66	49	20	335 070
Irbid	101	91	-0.05	76	49	19	917 820
Jarash	102	89	-0.30	51	49	18	149 850
Karak	98	86	-0.17	35	49	20	197 860
Ma'an	98	84	-0.16	44	48	18	89 010
Madaba	99	88	-0.07	60	48	19	129 890
Mafraq	97	83	-0.36	32	49	18	232 720
Tafileh	96	87	-0.13	65	49	18	73 740
Zarqa	101	92	-0.03	95	49	20	758 000

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.5. REGION-LEVEL DESCRIPTIVE STATISTICS FOR THE 1982 CENSUS OF MOROCCO

Region	Whipple index	Per cent literate	Wealth index	Per cent female	Median age	Population
Chaouia-Ouardigha	170	25	-0.28	49	17	1 307 800
Charb-Chrarda-Beni Hssen	183	28	-0.16	50	17	1 226 920
Doukala Abda	173	21	-0.25	50	18	1 476 060
Fès-Boulemane	164	35	0.10	50	18	947 060
Grand-Casablanca	115	55	0.82	50	20	2 362 260
Guelmin-Es-Samara	154	23	-0.34	53	17	242 420
Laayoune-Boujdour-Sakia El Hamra	98	30	-0.81	45	23	100 600
Marrakech-Tensift-Al Haouz	192	21	-0.19	50	18	2 269 860
Meknes-Tafilalet	166	31	-0.10	51	18	1 482 440
Oriental	137	31	0.08	51	17	1 414 300
Oued-Ed-Dahab-Lagouira	106	33	-0.82	47	23	12 820
Rabat-Salé-Zemmour-Zaer	146	44	0.36	50	19	1 420 300
Souss-Massa-Draa	176	20	-0.26	52	17	1 955 800
Tadla Azilal	181	21	-0.40	50	17	1 048 580
Tanger-Tétouan	175	32	0.03	50	18	1 449 780
Taza-Al Heiceima-Taounate	191	18	-0.58	51	16	1 436 420

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.6. REGION-LEVEL DESCRIPTIVE STATISTICS FOR THE 1994 CENSUS OF MOROCCO

Region	Whipple index	Per cent literate	Wealth index	Per cent female	Median age	Population
Chaouia-Ouardigha	121	38	-0.29	50	20	1 496 800
Charb-Chrarda-Beni Hssen	119	38	-0.25	50	19	1 620 900
Doukala Abda	128	32	-0.31	50	20	1 781 420
Fès-Boulemane	123	45	0.11	51	20	1 316 760
Grand-Casablanca	105	66	0.57	51	24	3 102 000
Guelmin-Es-Samara	112	39	-0.04	52	18	346 360
Laayoune-Boujdour-Sakia El Hamra	100	57	0.32	48	20	161 100
Marrakech-Tensift-Al Haouz	129	30	-0.26	50	20	2 706 700
Meknes-Tafilalet	126	43	-0.09	51	20	1 861 280
Oriental	111	43	0.16	51	21	1 714 960
Oued-Ed-Dahab-Lagouira	91	52	0.18	43	22	33 900
Rabat-Salé-Zemmour-Zaer	116	56	0.31	50	22	1 966 240
Souss-Massa-Draa	118	32	-0.24	51	19	2 618 740
Tadla Azilal	131	31	-0.42	50	19	1 322 700
Tanger-Tétouan	130	43	-0.07	50	19	2 016 780
Taza-Al Heiceima-Taounate	124	29	-0.62	51	18	1 722 380

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.7. REGION-LEVEL DESCRIPTIVE STATISTICS FOR THE 2004 CENSUS OF MOROCCO

Region	Whipple index	Per cent literate	Wealth index	Per cent female	Median age	Population
Chaouia-Ouardigha	114	49	-0.22	50	23	1 639 600
Charb-Chrarda-Beni Hssen	111	48	-0.16	50	22	1 844 780
Doukala Abda	118	42	-0.33	50	23	1 969 220
Fès-Boulemane	117	56	0.19	51	24	1 565 660
Grand-Casablanca	104	72	0.52	51	27	3 636 100
Guelmin-Es-Samara	107	54	0.04	52	22	419 540
Laayoune-Boujdour-Sakia El Hamra	106	66	0.45	48	23	242 540
Marrakech-Tensift-Al Haouz	112	42	-0.23	50	23	3 100 280

TABLE AII.7 (*continued*)

Region	Whipple index	Per cent literate	Wealth index	Per cent female	Median age	Population
Meknes-Tafilalet	117	53	-0.08	51	24	2 112 140
Oriental	111	53	0.06	51	25	1 876 420
Oued-Ed-Dahab-Lagouira	103	57	0.04	43	24	74 460
Rabat-Salé-Zemmour-Zaer	110	64	0.26	51	26	2 351 620
Souss-Massa-Draa	108	47	-0.15	52	23	3 083 140
Tadla Azilal	117	41	-0.45	52	23	1 439 120
Tanger-Tétouan	123	53	-0.07	50	22	2 440 540
Taza-Al Heiceima-Taounate	119	39	-0.63	51	21	1 789 340

Source: ESCWA calculations based on Minnesota Population Center, 2011.

TABLE AII.8. GOVERNORATE-LEVEL DESCRIPTIVE STATISTICS FOR THE 1997 CENSUS OF PALESTINE

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Per cent in camps	Population
Bethlehem	112	87	0.15	34	49	18	8	130 012
Deir al-Balah	115	87	0.00	31	50	15	66	145 464
Gaza	111	88	0.28	81	50	14	18	360 661
Hebron	108	84	-0.43	67	49	15	3	387 651
Jenin	102	86	-0.07	39	49	17	5	195 197
Jericho	94	83	-0.40	47	51	17	21	28 520
Jerusalem	111	89	0.25	40	49	17	6	111 655
Khan Yunis	109	86	-0.14	69	49	15	18	196 353
Nablus	104	88	0.12	41	49	18	11	251 067
North Gaza	120	86	0.14	63	49	13	34	180 137
Qalqilya	112	85	0.15	60	48	17	0	69 386
Rafah	116	85	-0.14	42	49	14	49	120 235
Ramallah and al-Bireh	104	86	0.12	34	50	17	6	201 929
Salfit	99	85	-0.02	28	50	17	0	46 456
Tubas	118	83	-0.54	34	49	17	13	34 551
Tulkarm	105	87	0.27	46	50	18	12	129 203

Source: ESCWA calculations based on data from the Palestinian Central Bureau of Statistics.

TABLE AII.9. GOVERNORATE-LEVEL DESCRIPTIVE STATISTICS FOR THE 2007 CENSUS OF PALESTINE

Governorate	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Per cent in camps	Population
Al Quds	108	95	-0.25	65	49	23	6	129 319
Ariha and Al Aghwar	97	92	-0.32	53	51	18	25	39 568
Beit Lahim	98	93	0.06	65	49	19	7	168 682
Deir Al balah	105	94	0.07	63	49	17	37	199 765
Gaza	101	95	0.16	91	49	16	7	482 422
Hebron	98	92	-0.13	86	49	16	3	538 515
Jenin	100	93	0.04	59	49	19	4	250 591
Khan Yunis	101	93	-0.06	80	49	17	14	267 943
Nablus	98	94	0.16	55	49	19	10	312 935
North Gaza	101	94	0.10	84	49	16	15	264 118
Qalqilia	106	93	0.16	62	50	18	0	87 540
Rafah	107	93	-0.13	76	49	17	20	169 401
Ramallah and al-Bireh	100	93	0.37	52	49	20	6	266 789
Salfit	96	92	0.16	37	49	18	0	59 233
Tubas	97	92	-0.04	68	49	19	12	47 086
Tulkarm	102	94	0.10	68	49	19	11	156 618

Source: ESCWA calculations based on data from the Palestinian Central Bureau of Statistics.

TABLE AII.10. REGION-LEVEL DESCRIPTIVE STATISTICS FOR THE 2008 CENSUS OF THE SUDAN

	Region	Whipple index	Per cent literate	Wealth index	Per cent urban	Per cent female	Median age	Population
	Central	240	62	0.99	23	52	18	7 309 510
	Darfur	269	41	0.24	19	50	15	7 107 417
	Eastern	254	46	0.57	31	48	18	4 388 123
	Khartoum	190	81	1.75	81	48	22	5 245 078
	Kordofan	260	41	0.40	21	52	16	4 231 931
	Northern	219	71	1.23	24	51	22	1 732 569
	Bahr El Ghazal	173	18	0.31	20	53	15	2 590 331
South	Equatoria	173	39	0.40	32	51	16	2 354 885
Sudan	Upper Nile	179	22	0.36	22	50	15	2 713 350

Source: ESCWA calculations based on Minnesota Population Center, 2011.

REFERENCES

Aickin, M. and H. Gensler (1996). Adjusting for multiple testing when reporting research results: the Bonferroni vs. Holm methods. *American Journal of Public Health*, vol. 86, No. 5 (May), pp. 726-728.

Alleaume, Ghislaine and Philippe Fargues (1998). La naissance d'une statistique d'État. *Histoire et Mesure*, vol. 13, No.1-2, pp. 147-193. Available from http://www.persee.fr/web/revues/home/prescript/article/hism_0982-.

Arriaga, Eduardo E. (1994). *Population Analysis with Microcomputers*. United States Census Bureau.

Courbage, Youssef and Philippe Fargues (1975). La population du Moyen Orient. *Population* (French edition), No. 6, November-December.

Courbage, Youssef and Philippe Fargues (1996). *Chrétiens et juifs dans l'Islam arabe et turc*. Paris: Petite Bibliothéque Payot.

Filmer, D. and L. Pritchett (1998). *Estimating Wealth Effect Without Expenditure Data–Or Tears: An Application to Educational Enrollments in States of India*. World Bank Policy Research Working Paper No. 1994. Washington, D.C.: World Bank.

Goldstein, Harvey (1999). *Multilevel Statistical Models.* London: Institute of Education, Multilevel Models Project.

Kolenikov, S. and G. Angeles (2009). Socioeconomic Status Measurement with Discrete Proxy Variables: is Principal Component Analysis a Reliable Answer? *Review of Income and Wealth*. March, Series 55, No. 1.

Minnesota Population Center (2011). Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis, Minnesota: University of Minnesota, 2011.

Mrabet, Achraf (2013). Presentation on the census of Tunisia at the ESCWA Expert Group Meeting on Evaluation of Census Data in Arab Countries, Beirut, 12-13 June.

Rutstein, Shea O. and Kiersten Johnson (2004). *The DHS Wealth Index.* DHS Comparative Reports No. 6. Calverton, Maryland: ORC Macro.

United Nations Statistics Division (2011). *2010 World Population and Housing Census Programme*. Sub-regional Workshop on Census Data Evaluation, Phnom Penh, 14-17 November.