

# **Measuring Adult Mortality Using Sibling Survival: A New Analytical Method and New Results for 38 Countries, 1970-2005**

Ziad Obermeyer

Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA  
Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, USA

Chang Park

Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA  
Mount Sinai School of Medicine, New York, NY, USA

Julie Knoll Rajaratnam

Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

Emmanuela Gakidou

Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

Alan D Lopez

School of Population Health, The University of Queensland, Brisbane, Queensland, Australia

Christopher JL Murray

Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

## **Abstract**

### *Background:*

For several decades, global public health efforts have focused on the development and application of disease control programs to improve child survival in developing populations. The need to reliably monitor the impact of such intervention programs in countries has led to significant advances in demographic methods and data sources, particularly with large-scale cross-national survey programs such as the Demographic and Health Surveys (DHS). While no comparable effort has been undertaken for adult mortality, the availability of large datasets with information on adult survival from censuses and household surveys offers an important opportunity to dramatically improve our knowledge about levels and trends in adult mortality in countries without good vital registration. To date, attempts to measure adult mortality from questions in censuses and surveys have generally led to implausibly low levels of adult mortality owing to biases inherent in survey data such as survival and recall bias. Recent methodological developments and the increasing availability of large surveys with information on sibling survival suggest that it may well be timely to reassess the pessimism that has prevailed around the use of sibling histories to measure adult mortality.

### *Methods and Findings:*

We present the Corrected Sibling Survival (CSS) method, which addresses both the survival and recall biases that have plagued the use of survey data to estimate adult mortality. Using logistic regression our method directly estimates the probability of dying in a given country, by age, sex and time period from sibling history data. The logistic regression framework borrows strength across surveys and time periods for the estimation of the age patterns of mortality, and facilitates the implementation of solutions for under-representation of high mortality families and recall bias. We apply the method to generate estimates of and trends in adult mortality, using the summary measure  ${}_{45}q_{15}$ , the probability of a 15 year old dying before his or her 60<sup>th</sup> birthday, for 38 countries with DHS sibling survival data.

Our findings suggest that levels of adult mortality prevailing in many developing countries are substantially higher than previously suggested. Generally, our estimates show the risk of adult death between ages 15 and 60 years to be about 20-30% for females and 30-40% for males in sub-Saharan African populations largely unaffected by HIV. In countries of Southern Africa, where the HIV epidemic has been most pronounced, 8 out of 10 men alive at age 15 will be dead by age 60, as will 6-7 out of 10 women. Adult mortality levels in populations of Asia and Latin America are generally lower than in Africa, particularly for women. The exceptions are Haiti and Cambodia where mortality risks are comparable to many countries in Africa. In all other countries with data, the probability of dying between ages 15 and 60 was typically around 10% for women and 20% for men, not much higher than the levels prevailing in several more-developed countries.

### *Conclusions:*

Our results represent a major expansion of direct knowledge of levels and trends in adult mortality in the developing world. The CSS method provides grounds for renewed optimism in collecting sibling survival data. We argue that all nationally representative survey programs ought to implement this critical module for tracking adult mortality. These opportunities must be seized if we are to more reliably understand the levels and patterns of adult mortality, and how they are changing.

## **Introduction**

For several decades, global public health efforts have focused on the development and application of disease control programs to improve child survival in developing countries. There are strong moral and pragmatic reasons for doing so. Technologies for preventing and successfully treating the leading causes of death in children are available, and their increasing effective use is leading to continuing declines in child mortality. The need to reliably monitor the impact of such intervention programs in countries has led to significant advances in demographic methods and data sources, particularly from large-scale cross national survey programs such as the Demographic and Health Surveys (DHS) [1-3]. As a consequence, levels and trends of child mortality are reasonably well understood in most countries [4,5].

No comparable effort has been undertaken for adult mortality, despite the impetus from the global HIV epidemic and the need to evaluate the progress of prevention strategies. At the same time, increasing global concern with avoiding premature adult death from tobacco and other established and modifiable causes of disease would be more effectively supported if progress with prevention strategies could be reliably monitored. We remain remarkably ignorant about current levels of adult mortality in the majority of developing countries, let alone how it is changing. Vital registration systems are generally incomplete in most developing countries; this has led to alternative measurement strategies for adult mortality [1-3]. Previous attempts to measure adult mortality from questions in censuses and surveys have generally led to implausibly low levels of adult mortality, leaving lingering doubts about the true impact of disease and injury on adult survival in developing countries [6,7]. Addressing this “scandal of ignorance” is becoming increasingly important as more and more children survive to adulthood as a consequence of improved child survival.

The availability of large datasets with information on adult survival from censuses and household surveys offers an important opportunity to dramatically improve our knowledge about levels and trends in adult mortality in countries without good vital registration, provided methods can be developed to adequately exploit them. In particular, questions about the survival of siblings of the respondent offer considerable promise to estimate mortality levels by age and sex at different periods of time prior to the survey. These ‘sibling histories’ theoretically can provide the information required to estimate mortality among adults. In practice, however, they suffer from a number of important limitations. Families which have disintegrated because of discord or death will be under-represented in population surveys, implying that mortality measures calculated from surveys will be biased downwards. It is also likely that respondents will fail to recall some deaths of siblings, especially those that occurred several years prior to the survey, or when the respondent had not had recent contact with his or her siblings. This latter concern is thought to be particularly problematic for reporting about the deaths of brothers, due to higher levels of male migration [6].

Recent methodological developments and the availability of new Demographic and Health Surveys containing information on sibling survival suggest that it may well be timely to reassess the pessimism that has prevailed around the use of sibling histories to measure adult mortality. Techniques to adjust for the underrepresentation of high mortality families in population surveys have been developed [8]. Building on this methodological advance, we have developed new methods to correct for the underestimation of mortality arising from the recall of deaths, as well as for the downward bias that arises in the estimation of recent levels of female mortality when the respondent is female with no reported siblings. We collectively refer to this new set of methods as the Corrected Sibling Survival (CSS) method. We have applied the CSS method to sibling survival data for 38 countries, primarily in Africa, where uncertainty about true levels and trends in adult mortality has been greatest.

While compilations of adult mortality estimates are routinely made by international organizations such as the United Nations [9], the World Health Organization [10] and the World Bank [11], these are largely model-based and rely heavily on indirect techniques and model life tables to infer mortality levels and patterns. There is considerably uncertainty about whether the demographic assumptions upon which these models are based are met in contemporary

developing populations. Most likely they are not, and hence mortality estimates based on them should be viewed cautiously. Concerns about data quality certainly affect the reliability of our results. However, while there remains uncertainty in the estimates we present in this paper, they nonetheless have the advantage of being directly calculated from actual country data.

## Methods

### *General Model*

Our method uses logistic regression to directly estimate the probability of dying for a country, by age, sex and time period from sibling history data. We transform survey respondents' recall of the dates of their siblings' births and deaths into a dataset where one observation indicates whether a sibling alive at the beginning of a given five-year period is alive or dead at the end of the period.

Survival analysis using logistic regression of reshaped data in this manner has a long history in epidemiology and demography [12-15]. The logistic regression framework borrows strength across surveys and time periods for the estimation of the age patterns of mortality and facilitates the implementation of solutions for under-representation of high mortality sibships<sup>1</sup> and the estimation of recall bias. Our model can be applied to a single country with multiple surveys, or any grouping of populations where at least some of them have had multiple surveys. The general model we estimate is:

$$\text{Logit}(Y_{ait} = 1) = \beta_0 + \beta I_a + \beta I_{it} + \beta(TiPS)$$

Where  $Y_{ait}$  indicates survival or death in age group  $a$ , in country  $i$  for a five-year period of time  $t$ ;  $I_a$  is a set of dummy indicators for each age group,  $a$ ;  $I_{it}$  is a set of dummy indicators for country  $i$  in the five year period designated by  $t$  and  $TiPS$  is a continuous variable representing the time prior to the survey, explained in detail below.

The classic work of Gompertz as well as more recent work shows that the age pattern of mortality follows a consistent pattern where the log of the death rate roughly increases linearly with age from young adult ages to the oldest ages [16,17]. By including a set of dummy variables for each age-group in the model, we borrow strength across all surveys to inform the age pattern of mortality without imposing a model-based structure *a priori*. It is clear, however, that certain epidemiological phenomena lead to different age-patterns of mortality, particularly those affecting young adults, such as HIV/AIDS. In addition, the age pattern of HIV-attributable mortality is likely to vary according to the level of sero-prevalence. The model is flexible and allows for different age patterns of mortality for sub-groups of country-years by incorporating a separate set of age-specific effects for those country-periods. This allows the model to empirically estimate the variation in the relative pattern of mortality by age.

Several previous studies have raised the possibility that respondents may omit reporting some sibling deaths [8,18,19]. To deal with the problem of recall bias, we include  $TiPS$ , a variable that

---

<sup>1</sup> We use the term sibship to refer to all children born to the same mother.

reflects the Time Prior to the Survey of the observation in question. For example, if a respondent to a survey carried out in the year 2000 reported on the death of a sibling in 1990, the value of *TiPS* would be 10 for that observation. The coefficient on the *TiPS* variable, when exponentiated, can be interpreted as the annual incremental reduction in the observed probability of death due to omitted death reports. This variable empirically measures recall bias and can be used to correct for it. Intuitively, *TiPS* captures the difference between deaths reported in the more recent periods of older surveys and deaths reported for more distant periods in more recent surveys. The coefficient can only be estimated if there is sufficient overlap of observations from different surveys for the same country-year. To estimate the level of adult mortality in the absence of recall bias, we set the coefficient on the *TiPS* variable to zero, and use the other model coefficients to generate estimates of  ${}_{45}q_{15}$  for each sex and for each country-time period.

Sibling-period observations are weighted to address the under-representation of high-mortality families in population-based surveys following the general methods proposed by [8]. This method (which we abbreviate GK) incorporates a family-level weight,  $W_f = B_f / S_f$ , where  $B_f$  is the original sibship size and  $S_f$  is the number of siblings in family  $f$  surviving to the time of the survey, which algebraically corrects for the under-representation of high-mortality families in the survey sample. Since our method directly estimates probability of death using individual-level data, we simply multiply  $W_f$  by the survey weight to generate the final weight for each observation.

Where, as is usual, the respondents to a sibling history survey are exclusively female, the sample by definition includes no sibships where 1) the siblings are all males, or 2) all the females have died. For both of these situations, the GK method does not adjust for the non-representation of these sibships. In the first case, mortality rates for males won't be affected if we assume that the mortality rates for males in all-male sibships are the same as mortality rates for males in sibships with one or more females (which are represented in all-female surveys). In the second case where all females in a sibship have died, however, mortality rates will be biased downward. In this case, female mortality rates in sibships with no female survivors (not reflected by the survey respondents) are *de facto* higher than mortality rates in sibships with one or more female survivors (which are reflected by the survey respondents). Therefore, we need to adjust the female rates to account for this bias. The reverse would be true if all respondents were male.

Figure 1 illustrates this zero-female-survivor bias using data from a DHS survey in Mali in which only females were interviewed. The figure shows the reported fraction of siblings who have died as a function of the initial number of males or females in the sibship. For males, there is a steady decline in the fraction dead as the initial number of male siblings declines; for females, there are no female deaths by definition where the initial number of female siblings was one but there is also an undercount of the fraction dead where the initial number of female siblings was two. This undercount is due to the fact that in some two female sibships both females have died. The undercount of percent dead continues but diminishes as the initial number of female siblings increases (because it becomes much less likely at greater sibship sizes that all females have died). A simple algebraic correction for the percent who have died can be estimated for female sibships of size two and three (see Appendix 1). We have used this formula to correct the percent dead for sibships of two and three females. Linear regression of female percent dead on the number of initial female siblings has been used to predict the true female percent dead for sibships with one

female. This corrected proportion is then used to estimate the number of missing female deaths from single-female sibships (these corrected numbers are also shown in Figure 1). These rates, which are also shown in Figure 1, are used to generate an overall estimate for the number of missing female deaths which are then applied to correct upwards the female  ${}_{45}q_{15}$  (see Appendix 1 for details).

We compute standard errors to allow for clustering at the primary sampling unit (PSU) level. In theory, the non-independence of the probability of death at the individual level (*i.e.*, the fact that a sibling's probability of death at time  $t$  is likely to be correlated with his probability of death at time  $t-1$ ) will lead to artificially low standard errors. This would typically be addressed by clustering standard errors at the individual level (equivalent to the concept of shared frailty in survival analysis); however, our approach of clustering errors at the higher PSU level produces even larger standard errors than clustering at the individual level, and we thus view it as a more conservative approach.

Uncertainty in the estimates of  ${}_{45}q_{15}$  is captured by taking 1000 simultaneous draws from the variance-covariance matrix of the model, producing 1000 estimates of  ${}_{5}q_x$  for every age group and 1000 corresponding estimates of  ${}_{45}q_{15}$ . The 25<sup>th</sup> and 976<sup>th</sup> ranked estimates of  ${}_{45}q_{15}$  define the 95% uncertainty interval. In each draw, we repeat the analytical steps detailed above so that uncertainty at each point is propagated into the final results.

#### *Application of the Model to 38 Countries Using the DHS*

We demonstrate the utility of this approach to directly measure adult mortality by taking advantage of the extensive dataset collected via the Demographic and Health Surveys (DHS). As of September 1, 2008, data were available from 194 surveys carried out in 75 countries since 1985. Each survey interviews randomly selected women, typically between the ages of 15 and 49, asking questions related to fertility, child mortality, and health behaviors. Of these, 79 surveys in 40 countries incorporate a sibling history (also known as the maternal mortality) module, which specifically asks respondents about the births and deaths of their siblings. Appendix Table 1 summarizes the characteristics of the sibling history modules from these surveys with various measures of survey quality. We have excluded two surveys from our analysis, and these are not included in Appendix Table 1. Based on published reports of poor data quality in the sibling history module, we exclude the Nigeria 1999 survey [20]. We also exclude Jordan 1997 because HIV sero-prevalence data (used as described below) were not available.

Since the DHS only interviews women up to 49 years of age, the data become increasingly sparse for the older age groups as we trace the sibling records back in time. To avoid small number problems, we only use sibling history data up to 15 years prior to the survey.

For the logistic regression model, we fix the age pattern of mortality for each five-year age group. We allow for the possibility that the age pattern of mortality varies in populations affected by HIV by including five distinct sets of age indicator variables, one for country-periods with mean lagged HIV sero-prevalence between zero and 1.9% (denoted by  $H_0$ ), and others for 2.0 to 6.9% ( $H_1$ ), 7.0 to 11.9% ( $H_3$ ), and 12.0% and over ( $H_4$ ), based on 2008 UNAIDS historical estimates [21]. We impose a 5-year lag on the sero-prevalence numbers to account for the lag between

HIV incidence and mortality [22]. It is important to note that including different sets of indicator variables does not impose a different mortality pattern on these country-periods, but rather allows the pattern to differ based on the empirical data. The only assumption we make is that HIV can affect the age pattern of mortality. We estimate the *TiPS* coefficient for all country-periods simultaneously yielding an average recall effect for males and females across the set of countries.

The final model, which is estimated separately for males and females, is:

$$\text{Logit}(Y_{ait} = 1) = \beta_0 + \beta_{1-8} I_a^{H0} + \beta_{9-16} I_a^{H1} + \beta_{17-24} I_a^{H2} + \beta_{25-32} I_a^{H3} + \beta_{33-40} I_a^{H4} + \beta_{41-196} I_{it}^{CY} + \beta_{197} \text{TiPS}$$

Where  $Y_{ait}$  indicates survival or death in age group  $a$ , in country  $i$  at time  $t$ ,  $I_{it}^{CY}$  is a set of dummy indicators for country  $i$  in the five year period designated by  $t$ . *TiPS* is a continuous variable representing time prior to the survey.  $I_a^{H0}$  through  $I_a^{H4}$  are five sets of dummy indicators for each five-year age group from 15 to 60 ( $a = 15-19, 20-24, \dots, 55-59$ ). Only one set of these age indicators is estimated, depending on the level of HIV sero-prevalence in that country and time period.

## Results

### *Corrected Sibling Survival model results*

The impact of the three components of the CSS model on the estimates of adult mortality rates is illustrated for six countries in Figure 2. Computation of  ${}_{45}q_{15}$  using the Gakidou-King weights has a major effect on measured levels of adult mortality for males and females in all countries. This indicates that under-representation of high mortality sibships is an important aspect when analyzing sibling survival data. On average, the GK weights raise the estimated  ${}_{45}q_{15}$  by 27% ranging from 5 to 55% across country-periods.

Regression results for the *TiPS* variable are summarized in Table 1. For males, the coefficient is -.0299, representing a 2.9% decrease in reported deaths for each additional year prior to the survey. For females, the annual decrease is lower at 1.9%. We have also estimated the value of *TiPS* coefficient separately for each country, for the six countries where three or more surveys with sibling history modules are available. These results are summarized in Figure 3 and show that the estimated decline in deaths reported due to recall bias varied from -1.0% (Mali females) to 8.6% (Madagascar males) across this set of countries. These values do not differ significantly from the mean effect across all countries except for females in Madagascar, Uganda, and Mali and for males in Madagascar and Uganda. Figure 2 illustrates how inclusion of the *TiPS* variable in the model leads to higher estimates of adult mortality especially in time periods further removed from the survey year.

Corrections for sibships where all females have died lead to much more modest changes in the estimated rates of adult female mortality. Table 2 summarizes the magnitude of these corrections for each country ranging from 1.0% to 4.5%. While conceptually important, these corrections do not profoundly alter the estimated levels of adult mortality. Overall, the combined effect of the

GK weights, the *TiPS* recall bias correction and the correction for missing female sibships leads to a profound increase in adult mortality rates estimated from sibling histories.

### *Levels of Mortality in 38 Countries*

Figure 4 shows CSS estimates of adult mortality for select countries; the same graphs for all countries included in our study can be found in the Appendix Figure. Our findings suggest that levels of adult mortality prevailing in many developing countries are substantially higher than previously suggested [7,23,24]. In sub-Saharan African populations largely unaffected by HIV, we estimate the risk of death between ages 15 and 60 years to be 20-30% for females and 30-40% for males, though considerable heterogeneity exists among countries. In Southern African countries where the HIV epidemic has been most pronounced, rates are uniformly and strikingly high: at current rates, 8 out of 10 men alive at age 15 will be dead by age 60, as will 6 to 7 out of 10 women. At the height of the Rwandan genocide in 1994, the probability of death between 15 and 60 based on prevailing mortality rates was close to 100%, but has since declined to levels more typical of sub-Saharan Africa. Mortality rates appear to be lower in West African populations for which data are available, notably Senegal, where the risk of adult death is around 20-25%.

Adult mortality levels in populations of Asia and Latin America are generally lower than in Africa, particularly for women, though Haiti and Cambodia are notable exceptions where mortality risks are comparable to many countries in Africa. In all other developing countries with available sibling history data, the probability of dying between ages 15 and 60 was typically around 10% for women and 20% for men; this is not much higher than estimated levels for many more developed countries [24].

### *Comparison of Corrected Sibling Survival to Other Measurements of Mortality*

Few countries with complete vital registration systems have included sibling survival data in national surveys. Validation of CSS thus depends on comparisons in those few countries which have vital registration data, demographic surveillance sites and deaths in the last 12 months collected in national censuses. Even though all of these are likely problematic comparators, Figure 4 compares our estimates to these three types of available data.

While vital registration is typically considered the gold standard for measurement of mortality, data from Peru, Brazil, Guatemala, Dominican Republic, Philippines, and South Africa are likely to be undercounts of national adult mortality rates. Data from the Dominican Republic and Peru are thought to be particularly incomplete, missing about 50% of adult deaths, while routine systems appear to be capturing the majority of deaths in countries such as Brazil, Guatemala, the Philippines and South Africa [25]. Figure 4 shows that in Guatemala, CSS estimates and vital registration data are similar for most years, although the CSS captures higher levels of adult mortality in the early 1980s, coincident with the outbreak and intensification of Shining Path guerilla warfare. A similar level of concordance is seen for the Philippines. In Brazil, CSS results for females appear to be somewhat lower than for the vital registration, but comparable overall. In Dominican Republic the vital registration rates appear to be implausibly low, especially for males, while the CSS presents more realistic levels. Overall, the levels of mortality suggested by



application of our methods are comparable or higher than what is suggested from vital registration systems, which are known to generally undercount deaths in developing countries.

In some countries, Demographic Surveillance Sites (DSS) have been operating that capture vital events that occur in defined populations. While these sentinel sites are quite small (covering populations of 30,000 to one million), and are typically selected expediently rather than randomly, they nonetheless can be a useful source of information on mortality and fertility levels. Where these sites are operative in the countries in our dataset, Figure 4 also shows the implied levels of adult mortality compared with CSS estimates. In most cases (e.g. South Africa, Tanzania, Senegal, Mozambique, Burkina Faso), DSS death rates for most time periods fall within the range of uncertainty suggested by our methods.

Finally, some national censuses collect data on deaths within households. As with vital registration data, censuses have varying levels of completeness, and like sibling history data, household deaths reported in censuses may be underreported [26]; however, in general, mortality data from well-implemented censuses are considered by demographers to be of good quality. Figure 4 shows that census-based mortality rates in Zimbabwe, Malawi, and Tanzania are remarkably close to the CSS estimates at various time periods, whereas in Mauritania and Ethiopia census data yield dramatically lower levels.

While this is not necessarily validation, the fact that CSS yields estimates that are comparable to those from independent data collection schemes is reassuring.

### *Mortality Trends*

Figure 4 also shows how risks of adult death have changed over the period 1980-2005, encompassing the peak effects of the HIV/AIDS epidemic, particularly in Africa. In some countries, notably Cote d'Ivoire, Cameroon, Kenya, Lesotho, Malawi, and Swaziland, death rates among adults appear to have risen throughout the past two decades or so. In Malawi and Zimbabwe, they have increased 3 to 4 fold since the late 1980's, with CSS showing the full devastation of the epidemic on adult survival. In Kenya, Zambia, Swaziland and Tanzania, death rates have doubled in 20 years, although in Tanzania at least, there are signs that death rates may be stabilizing. In others, particularly Benin, Congo, Ethiopia, and Madagascar, we have identified increases in mortality followed by declines. The effect of the 1994 genocide in Rwanda can be clearly seen, after which death rates dropped to levels similar to neighboring African countries. Haiti, Morocco, Peru and Senegal have experienced consistent declines in adult mortality over the past two decades. While it is difficult to interpret short term changes in death rates, except for the genocide in Rwanda, the utility of the method in determining longer term trends in mortality levels is clearly of great public health importance.

A summary appraisal of trends in adult mortality in Africa can be obtained from Figure 5 which shows levels of mortality estimated for various countries around 1990 (1988-1992) and 2000 (1998-2002). For women, the dramatic rise in adult mortality in Zambia is clear, as is the relatively low levels of mortality prevailing around 1990 in countries such as Morocco, Senegal, Benin, Kenya and South Africa. The greater heterogeneity of mortality levels around 2000 can also be seen, largely due to the differential impact of the HIV epidemic. For men, the

heterogeneity among countries is even greater, even in 1990, potentially reflecting the greater risks they incur from injuries and violence. The impact of the HIV epidemic is also clear from Figure 5, particularly in the Southern African states around 2000.

## **Discussion**

In this paper, we present an improved method for analyzing sibling survival data and demonstrate its application using 77 surveys undertaken in 38 countries. This represents a major expansion of our direct knowledge of levels and trends in adult mortality in developing countries. Expanded direct measurement will decrease the dependence of the global health community on uncertain predictions of adult mortality from levels of child mortality for tracking progress towards major health and development targets.

Collective concerns about the low levels of adult mortality from crude analysis of sibling data [1,18] may have dampened enthusiasm for collecting this type of data. We believe that the CSS method provides grounds for renewed optimism in collecting sibling survival data. Our experience strongly argues for all DHS surveys to incorporate the sibling history module, as even in middle-income countries, this information could be a useful adjunct to analyzing levels and trends in adult mortality from vital registration data, especially, as is likely, where it may only be partially complete.

While the sibling history module can be time consuming, its utility for public health suggests that all health surveys might include this module. Other multi-country survey programs such as the UNICEF Multiple-Indicator Cluster Survey (MICS) and the World Bank Living Standards Measurement Survey (LSMS) could well incorporate this approach. Widespread collection of these data will greatly strengthen our capacity to monitor maternal mortality and the ultimate effect of interventions such as anti-retrovirals in reducing adult mortality.

In addition to expanding the collection of sibling survival data in more surveys, our analysis suggests that the set of respondents who answer the sibling module in a survey should be expanded. By only asking women aged 15-49, the current DHS practice limits our ability to effectively measure mortality in adults over age 50 and for older time periods, especially for more than 15 years prior to the survey. Further, if both male and female respondents were to be asked the sibling history module, exploration of sex-specific biases in the recall of births and deaths of sisters and brothers would be possible and would allow for cross-validation. There is some evidence from our analysis that this effect may be significant. With increased concerns about the early impact of the epidemiological transition in many developing countries, expanding the age range of respondents will allow direct measurement of middle-aged mortality in these countries.

Given that survey teams in the DHS and other survey programs are already visiting households, expanding the set of respondents who are asked the sibling history module would not imply a substantial marginal cost on survey implementers. Moreover, our experience suggests that the information obtained is likely to be well worth the investment.

A key limitation of our analysis has been the estimation of the average recall bias across all surveys and the use of this average effect in calculating levels and trends in *45q15*. As more countries accumulate three or more surveys, it will be possible to apply the CSS model on a country by country basis. Country-specific recall bias parameter values can then be used in generating country-specific levels of adult mortality. As more countries collect sibling survival data, it will hopefully be possible to explore the contextual, linguistic and other cultural factors that might account for variability in recall bias. This type of qualitative insight should help to guide further improvements in survey instruments for sibling recall.

Measurement of adult mortality using sibling survival data will benefit from research on the wording and ordering of the sibling history module as well as from further statistical model refinement. Item wording, ordering and the use of prompters may aid recall of both living and dead siblings. Experiments are underway using alternative wording in Tanzania, India and the Philippines as part of the Gates Grand Challenges in Global Health initiative [27]. Comparison of the DHS sibling history data to UN Population Division data shows discrepancies between the average number of siblings by age of respondent and the corresponding total fertility rate at the time that the respondent was born. Further research is needed to better understand under what circumstances sibling recall of fertility and exogenous estimates of fertility differ. Similar investigations are needed to understand variations in the sex ratio of reported living and dead siblings. Growing recognition of the potential utility of sibling history data for public health monitoring will hopefully stimulate more research in this area.

The prospect that robust information on the levels and trends in adult mortality can be derived from periodic household surveys in low-income countries may warrant a reconsideration of the priorities for improved assessment of adult mortality. The MOVE group [28] called for an expansion of vital registration systems and the use of sample registration systems in the interim. Those are important initiatives but our findings suggest that it may be as important to more persuasively argue for the inclusion of sibling survival modules in ongoing survey programs. Further work is also needed to explore the feasibility of using new verbal autopsy instruments and analytical methods [29,30] in conjunction with these modules to ascertain not only death rates, but also causes of death. The demand for accountability and the use of pay for performance investments such as GAVI is likely to increase the pressure on countries to mount more frequent household surveys. Maximum use of these opportunities should be made for tracking trends in adult mortality.

These opportunities must be seized if we are to more reliably understand the levels, patterns, and causes of adult mortality, and how they are changing. Parallel investments in vital registration systems, and the routine inclusion of sibling survival questions into existing or planned household survey programs, are urgently needed if we are to rapidly build the evidence base for public health action. The success of child survival programs, accompanied by greater global concern for controlling major threats to health, argue for much greater research attention to be given to measuring adult mortality, and its causes. Keeping children alive to adulthood is a noble and worthy aim for the global public health community; keeping young adults alive and in good health until they reach old age should be seen as just as important.

## Acknowledgements

The authors gratefully acknowledge Mie Inoue and Danzhen You for assistance in obtaining census data, Stanislava Nikolova for assistance with referencing, and Margaret Hogan and other research staff at the Institute for Health Metrics and Evaluation for valuable comments on the research.

**Funding sources:** This study was funded in part by a grant from the Bill and Melinda Gates Foundation to the Institute for Health Metrics and Evaluation in Seattle, Washington, USA. The funding organization had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

**Author contributions:** All authors contributed to the writing of the manuscript and provided input into the analysis.

**Competing interests:** The authors declare that no competing interests exist.

## References

1. Gakidou E, Hogan M, Lopez AD (2004). Adult mortality: time for a reappraisal. *Int J Epidemiol* 33(4): 710-717.
2. Hill K, Lopez AD, Shibuya K, Jha P (2007). Interim measures for meeting needs for health sector data: births, deaths, and causes of death. *Lancet* 370: 1726-1735.
3. Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon FC, Rao C et al. (2007). Civil registration systems and vital statistics: successes and missed opportunities. *Lancet* 370: 1653-1663.
4. Murray CJ, Laakso T, Shibuya K, Hill K, Lopez AD (2007). Can we achieve Millennium Development Goal 4? New analysis of country trends and forecasts of under-5 mortality to 2015. *Lancet* 370(9592): 1040-1054.
5. The United Nations Children's Fund (UNICEF) (2006). *The state of the world's children 2007: women and children - the double dividend of gender equality*. New York: UNICEF.
6. Bicego G (1997). Estimating adult mortality rates in the context of the AIDS epidemic in sub-Saharan Africa: analysis of DHS sibling histories. *Health Transition Review* 7(S2): 7-22.
7. Timaeus IM, Jasseh M (2004). Adult mortality in sub-Saharan Africa: evidence from Demographic and Health Surveys. *Demography* 41(4): 757-772.

8. Gakidou E, King G (2006). Death by survey: estimating adult mortality without selection bias from sibling survival data. *Demography* 43(3):569-585.
9. United Nations Statistics Division (2008). *Demographic yearbook*. New York: United Nations. <http://unstats.un.org/unsd/demographic/products/dyb/dyb2.htm>.
10. World Health Organization (2008). *World Health Statistics 2008*. Geneva: World Health Organization. <http://www.who.int/whosis/whostat/2008/en/index.html>.
11. World Bank. HNPStats database. Washington, DC: World Bank. <http://go.worldbank.org/VK3WB23I40>.
12. Abbott RD (1985). Logistic regression in survival analysis. *American Journal of Epidemiology* 121(3): 465-471.
13. Adebayo SB, Fahrmeir L (2005). Analysing child mortality in Nigeria with geoaddivitive discrete-time survival models. *Stat Med* 24(5): 709-728.
14. Efron B (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association* 83: 414-425.
15. Kalbfleisch JD, Prentice RL (1973). Marginal likelihoods based on Cox's regression and life model. *Biometrika* 60(2): 267-278.
16. Gompertz B (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* 115: 513-583.
17. Hill K. Methods for measuring adult mortality in developing countries: a comparative review. *The Global Burden of Disease 2000 in Aging Populations Research Paper No. 01 .13*. Cambridge, MA: Harvard Burden of Disease Unit. <http://www.hsph.harvard.edu/burdenofdisease/publications/papers/Methods%20for%20Measuring%20Adult%20Mortality.pdf>.
18. Stanton C, Abderrahim N, Hill K (2000). An assessment of DHS maternal mortality indicators. *Stud Fam Plann* 31(2): 111-123.
19. Timaeus IM (1991). Measurement of adult mortality in less developed countries: a comparative review. *Popul Index* 57(4): 552-568.
20. National Population Commission. Maternal and child health. In: *Nigeria Demographic and Health Survey 1999*. Calverton (Maryland): National Population Commission and ORC/Macro. <http://www.measuredhs.com/pubs/pdf/FR115/08Chapter08.pdf>.
21. UNAIDS (2008). *Report on the global HIV/AIDS epidemic 2008*. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS).

22. Zaba B, Whiteside A, Boerma JT (2004). Demographic and socioeconomic impact of AIDS: taking stock of the empirical evidence. *AIDS* 18 Suppl 2: S1-S7.
23. United Nations (2008). Demographic Yearbook 2006. United Nations Statistics Division. <http://unstats.un.org/unsd/demographic/products/dyb/dyb2006.htm>.
24. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL, eds. (2006). Global burden of disease and risk factors. Oxford University Press and the World Bank.
25. Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD (2005). Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bull World Health Organ* 83(3): 171-177.
26. Banister J, Hill K (2004). Mortality in China 1964-2000. *Popul Stud (Camb)* 58(1): 55-75.
27. Walgate R (2003). Gates Foundation picks 14 grand challenges for global disease research. *Bull World Health Organ* 81(12): 915-916.
28. Lopez AD, AbouZahr C, Shibuya K, Gollogly L (2007). Keeping count: births, deaths, and causes of death. *Lancet* 370(9601): 1744-1746.
29. Murray CJ, Lopez AD, Feehan DM, Peter ST, Yang G (2007). Validation of the symptom pattern method for analyzing verbal autopsy data. *PLoS Med* 4(11): e327.
30. Obermeyer Z, Murray CJ, Gakidou E (2008). Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from the world health survey programme. *BMJ* 336(7659): 1482-1486.

## Appendix 1: Derivation of zero-female-survivor correction

In order to calculate the correction factor to account for the lack of representation of families with no female survivors, we employed the following procedure. We can compute algebraically the undercount of female deaths for sibships of size  $k$ :

$$D_k^{obs} = \pi \cdot N_k \cdot k - \pi^k \cdot N_k \cdot k$$

where  $D_k^{obs}$  is the number of female deaths observed in sibships of size  $k$  females,  $\pi$  is the true percent females dead in sibships of size  $k$ , and  $N_k$  is the total number of sibships of size  $k$  females. Therefore, in the equation above,  $\pi \cdot N_k \cdot k$  is the true number of deaths in sibships of size  $k$ , and  $\pi^k \cdot N_k \cdot k$  is the number of female deaths in sibships of size  $k$  females which are not observed because these deaths occur in sibships where all females die.

The observed percent of females who have died in sibships of size  $k$  females is represented by  $\pi_{obs}$  and is equal to the number of deaths observed in sibships of size  $k$  divided by the total number of females observed in sibships of size  $k$  ( $N_k^{obs}$  is the total number of sibships of size  $k$  females which we observe):

$$\pi_{obs} = \frac{D_k^{obs}}{N_k^{obs} * k}$$

Denoting  $N_k$  as the total number of sibships of size  $k$  females, then  $N_k^{obs} = N_k - \pi^k * N_k$ .

By substituting for  $D_k^{obs}$  and  $N_k^{obs}$ , we then arrive at an expression equating the observed percent dead with the true percent dead:

$$\pi_{obs} = \frac{D_k^{obs}}{N_k^{obs} * k} = \frac{\pi * N_k * k - \pi^k * N_k * k}{(N_k - \pi^k * N_k) * k} = \frac{(\pi - \pi^k) N_k * k}{(1 - \pi^k) N_k * k} = \frac{(\pi - \pi^k)}{(1 - \pi^k)}$$

While the mathematical solutions for  $\pi$  are infinite if  $k = 1$ , they are relatively easy to compute if  $k = 2$  or  $3$ . When  $k > 3$ , the solutions become quite complex. Fortunately, as sibship size increases, the contribution to total sibships of size  $k$  females of sibships where all females have died becomes substantively insignificant. Therefore, we use the algebraic solutions to correct the percent dead for sibship sizes 2 and 3 and then assume a linear relationship between (corrected) percent dead and sibship size to predict the true percent dead for sibships of only 1 female. Using these corrected percentages and the observed number of deaths for each sibship size, we compute the percent of total female deaths which are not captured due to the zero-surviving-female phenomenon. Finally, we correct our estimates of  ${}_{45}q_{15}$  upward by this factor.

## Tables

Country	Females		Males	
	coefficient	annual decline	coefficient	annual decline
Indonesia	-0.0201	1.99%	-0.0335	3.30%
Madagascar	-0.0618	5.99%	-0.0897	8.58%
Malawi	-0.0192	1.90%	-0.0337	3.31%
Mali	0.0101	-1.01%	-0.0109	1.09%
Peru	-0.0315	3.10%	-0.0371	3.64%
Uganda	-0.0007	0.07%	-0.0052	0.52%
Zimbabwe	-0.0216	2.13%	-0.0270	2.66%
All countries	-0.0196	1.94%	-0.0299	2.94%

**Table 1: TiPS regression coefficients and annual decline in mortality rates per year prior to the survey attributable to recall bias**



<b>Country</b>	<b>Zero-female survivor correction</b>
Benin	2.19%
Bolivia	2.63%
Brazil	1.32%
Burkina Faso	2.83%
Cambodia	3.19%
Cameroon	2.67%
Central African Republic	4.48%
Chad	4.05%
Congo	2.41%
Côte d'Ivoire	2.23%
Dominican Republic	1.95%
Eritrea	4.32%
Ethiopia	3.66%
Gabon	2.05%
Guatemala	2.01%
Guinea	4.13%
Haiti	2.58%
Indonesia	2.57%
Kenya	1.21%
Lesotho	2.65%
Madagascar	2.32%
Malawi	2.54%
Mali	3.24%
Mauritania	4.08%
Morocco	1.39%
Mozambique	4.06%
Namibia	2.12%
Niger	3.16%
Peru	1.33%
Philippines	1.00%
Rwanda	4.05%
Senegal	2.98%
South Africa	2.94%
Swaziland	2.42%
Tanzania	2.29%
Togo	2.14%
Uganda	3.00%
Zambia	2.45%
Zimbabwe	1.47%

**Table 2: Percent of deaths missing due to zero-female-survivors and the factor by which estimates are corrected upward as a result**

Country	Year	Total respondents to sibling history module	Total siblings reported	Alive/dead status unknown (percent of total siblings)	Total deaths	Sex unknown (percent of alive siblings)	Sex unknown (percent of total deaths)
Benin	1995	5488	36414	0.29%	8060	0.28%	2.39%
Benin	2005	17353	114323	0.16%	18508	0.13%	0.46%
Bolivia	1992	8603	52542	0.12%	8003	0.00%	1.11%
Bolivia	2002	17251	114508	0.44%	17952	0.03%	1.07%
Brazil	1995	12574	91122	2.73%	12356	0.02%	6.85%
Burkina Faso	1997	6425	41095	0.83%	7756	0.11%	4.46%
Burkina Faso	2003	12219	78102	0.26%	12746	0.06%	2.14%
Central African Republic	1993	5884	36328	0.04%	5788	0.00%	0.00%
Cambodia	1999	15351	94203	0.96%	16156	0.06%	0.06%
Cambodia	2004	16517	106463	0.73%	18910	0.00%	0.00%
Cameroon	1997	5489	38552	0.71%	6624	0.58%	3.20%
Cameroon	2003	10656	76311	0.05%	14433	0.00%	0.06%
Chad	1995	7448	48803	0.17%	9826	0.06%	1.01%
Chad	2003	6085	41885	0.00%	8827	0.00%	0.00%
Congo	2004	7051	47800	0.08%	7593	0.00%	0.01%
Cote d'Ivoire	1993	8099	52386	0.01%	7704	0.00%	0.00%
Cote d'Ivoire	2004	4741	30572	0.53%	4811	0.04%	0.69%
Dominican Republic	2001	11458	77542	0.78%	7297	0.05%	1.41%
Dominican Republic	2007	27153	168399	0.87%	14829	0.04%	0.83%
Eritrea	1994	5054	32137	0.14%	6359	0.12%	0.25%
Ethiopia	1999	15344	103915	0.59%	25457	0.02%	0.46%
Ethiopia	2004	13674	90386	0.58%	16896	0.12%	0.14%
Gabon	1999	6183	41507	0.01%	5600	0.24%	1.38%

**Appendix Table 1: Characteristics of the sibling history surveys in 77 DHS**

Country	Year	Total respondents to sibling history module	Total siblings reported	Alive/dead status unknown (percent of total siblings)	Total deaths	Sex unknown (percent of alive siblings)	Sex unknown (percent of total deaths)
Guatemala	1994	12367	85677	0.40%	12538	0.01%	0.44%
Guinea	1998	6753	39417	0.19%	7721	0.06%	0.12%
Guinea	2004	7613	48771	0.12%	10955	0.00%	0.03%
Haiti	1999	10159	70025	0.11%	15105	0.03%	0.21%
Haiti	2004	10527	73190	0.05%	14847	0.00%	0.00%
Indonesia	1993	28168	162020	0.09%	21122	0.04%	1.23%
Indonesia	1996	28810	161095	0.05%	16448	0.00%	0.38%
Indonesia	2001	28121	161137	0.11%	14738	0.01%	0.66%
Kenya	1997	7873	57618	0.26%	6241	0.07%	1.41%
Kenya	2002	8176	59811	0.22%	7584	0.04%	1.08%
Lesotho	2003	6921	41359	0.11%	6868	0.02%	0.39%
Madagascar	1991	6260	46332	0.10%	6696	0.00%	0.00%
Madagascar	1996	7060	51115	0.45%	5928	0.04%	1.55%
Madagascar	2002	7644	50647	0.19%	3615	0.16%	0.55%
Malawi	1991	4849	34585	0.46%	8339	0.19%	0.84%
Malawi	2000	13220	91992	0.03%	21699	0.03%	0.30%
Malawi	2003	11356	74818	0.04%	14383	0.03%	0.15%
Mali	1994	9704	63108	0.01%	14154	0.00%	0.10%
Mali	2000	12814	82971	0.43%	16576	0.10%	0.38%
Mali	2005	14050	98150	0.36%	21383	0.04%	0.25%
Mauritania	1999	7728	49517	0.21%	5887	0.05%	0.07%

**Appendix Table 1: Characteristics of the sibling history surveys in 77 DHS**

Country	Year	Total respondents to sibling history module	Total siblings reported	Alive/dead status unknown (percent of total siblings)	Total deaths	Sex unknown (percent of alive siblings)	Sex unknown (percent of total deaths)
Morocco	1991	9256	69757	0.07%	10665	0.00%	0.01%
Morocco	2002	16606	125853	0.07%	18249	0.00%	0.25%
Mozambique	1996	8732	50201	0.01%	7276	0.00%	0.00%
Mozambique	2002	12003	76087	0.35%	13243	0.03%	0.47%
Namibia	1991	5421	36848	0.68%	4437	0.12%	1.26%
Namibia	1999	6750	43609	0.33%	4597	0.16%	0.54%
Nepal	1996	8429	52580	0.15%	11666	0.00%	0.06%
Nepal	2006	10653	65060	0.08%	11610	0.00%	0.08%
Niger	1991	6503	43950	0.12%	9815	0.02%	0.78%
Niger	2005	8976	64823	0.12%	12616	0.01%	0.05%
Peru	1990	15882	101628	0.17%	10775	0.00%	0.00%
Peru	1995	28951	201368	0.73%	29717	0.03%	2.21%
Peru	1999	27779	188782	0.58%	27293	0.01%	1.66%
Peru	2002	12217	81688	0.10%	10197	0.02%	0.16%
Philippines	1992	15029	103636	0.26%	8404	0.03%	0.24%
Philippines	1997	13977	94979	0.16%	9023	0.02%	0.27%
Rwanda	1999	10415	75667	1.81%	19665	0.10%	0.46%
Rwanda	2004	11181	83574	1.06%	22441	0.01%	0.06%
Senegal	1991	6310	41913	0.09%	8160	0.01%	0.27%
Senegal	2004	14371	100597	0.16%	15674	0.02%	0.54%

**Appendix Table 1: Characteristics of the sibling history surveys in 77 DHS**

Country	Year	Total respondents to sibling history module	Total siblings reported	Alive/dead status unknown (percent of total siblings)	Total deaths	Sex unknown (percent of alive siblings)	Sex unknown (percent of total deaths)
SouthAfrica	1997	11717	63482	1.02%	6319	0.32%	2.20%
Swaziland	2005	4806	30555	0.37%	4001	0.03%	0.42%
Tanzania	1995	8118	55738	0.09%	8268	0.08%	0.36%
Tanzania	2003	10203	75925	0.08%	12990	0.01%	0.11%
Togo	1997	8569	58045	0.11%	11410	0.12%	0.65%
Uganda	1994	7066	51375	0.40%	9375	0.08%	1.06%
Uganda	1999	7239	54233	0.94%	10493	0.11%	0.81%
Uganda	2005	8519	65904	0.97%	15062	0.01%	0.21%
Zambia	1995	8021	58422	0.05%	9809	0.04%	0.13%
Zambia	2000	7658	55459	0.10%	9828	0.02%	0.06%
Zimbabwe	1993	6128	44358	0.08%	5582	0.04%	0.59%
Zimbabwe	1998	5907	40441	0.23%	4437	0.02%	0.50%
Zimbabwe	2004	8651	55714	0.10%	6932	0.09%	0.13%

**Appendix Table 1: Characteristics of the sibling history surveys in 77 DHS**

## **Figure Legends:**

**Figure 1: The relationship between percent dead and sibship size for males and females, using an example from the Mali 2005 DHS.** The percent dead for females has been corrected in sibship sizes 1, 2, and 3 to account for the increased occurrence of zero-female-survivors in these sibships.

**Figure 2: A step-by-step look at each of the adjustments in the Corrected Sibling Survival (CSS) Method:** the effects on mortality estimates of the Gakidou-King survival bias adjustment, adjusting for recall, and the zero-female-survivor correction.

**Figure 3: Country-specific estimates of recall bias** as measured by the TiPS coefficient compared to the overall estimate from CSS model (red line)

**Figure 4: Estimates of 45q15 from the CSS method compared to estimates generated from vital registration, demographic surveillance sites, and census household death estimates** in countries for which these comparison estimates are available.

**Figure 5: CSS-generated estimates of 45q15 for African countries with DHS data, 1990 and 2000**

**Appendix Figure: Estimates of 45q15 from the CSS method for all 38 countries where sibling history surveys are available in the Demographic and Health Surveys**

Figure 1

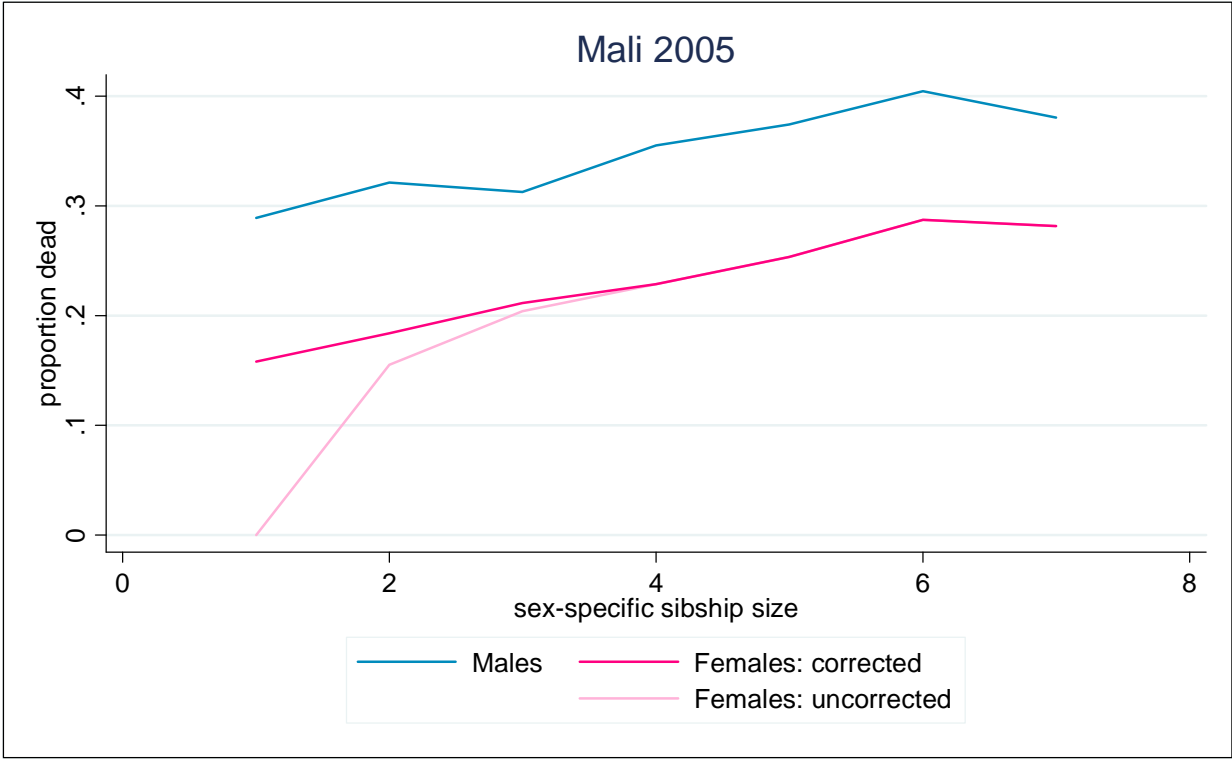


Figure 2

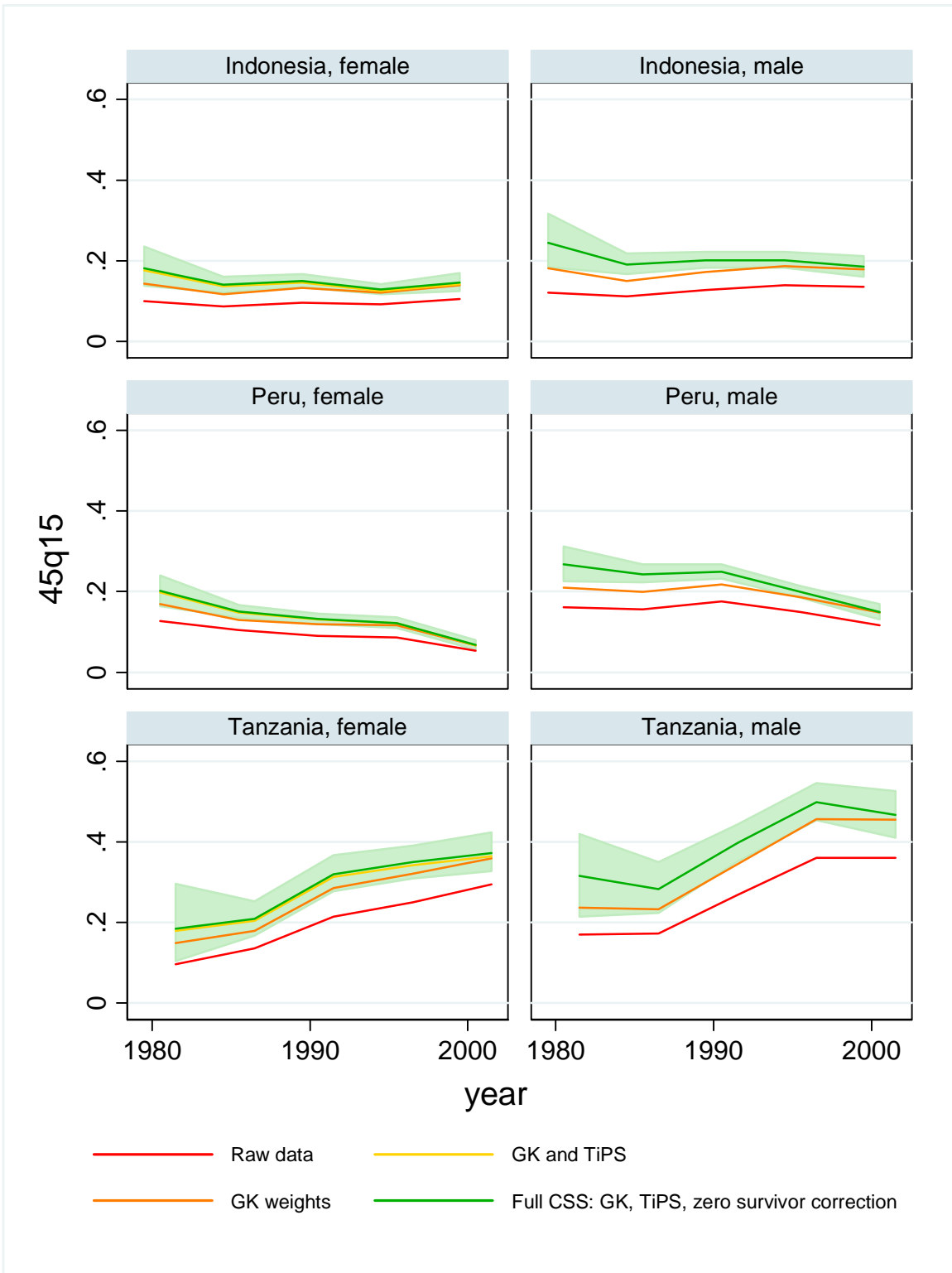




Figure 2

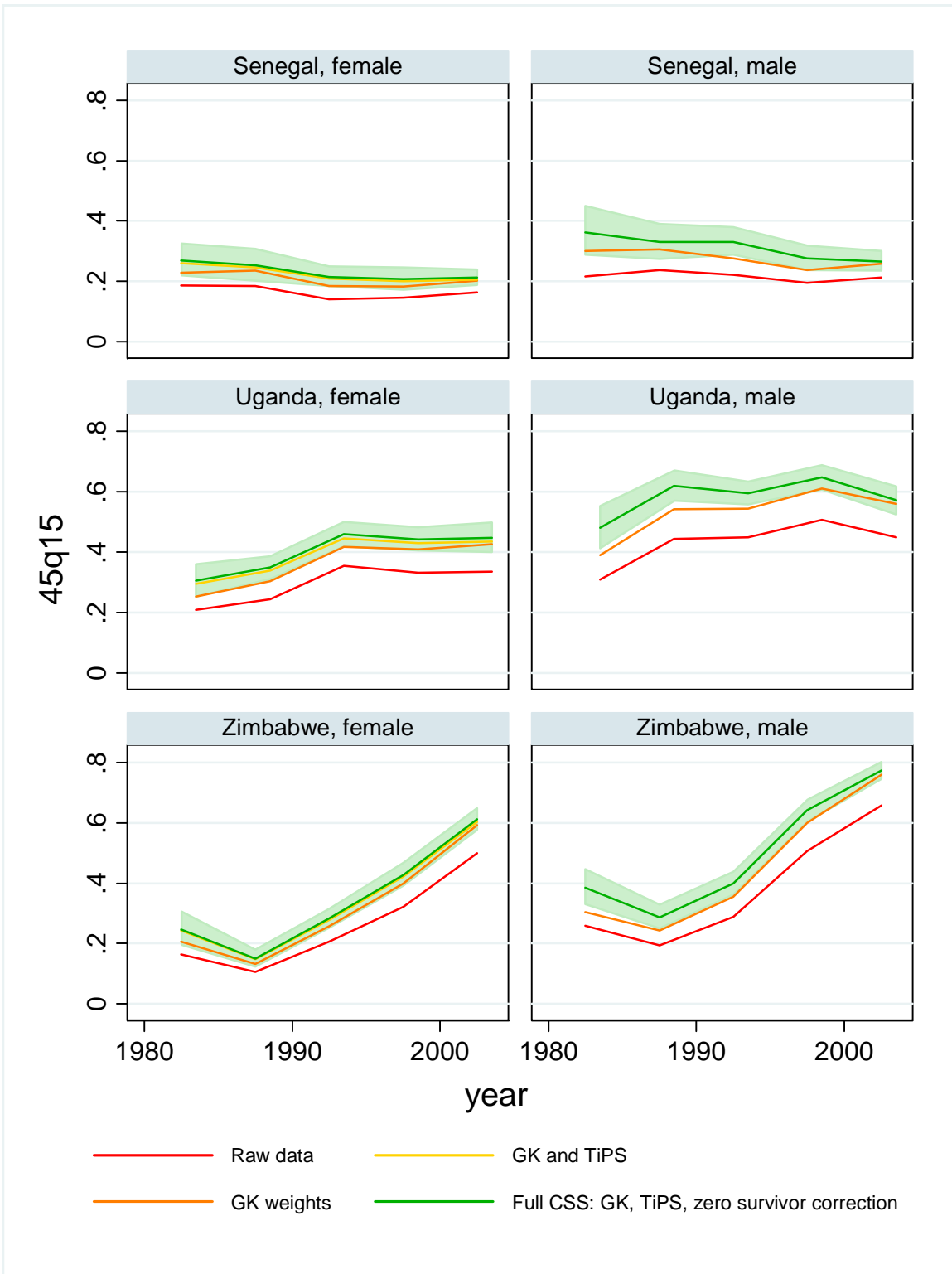


Figure 3

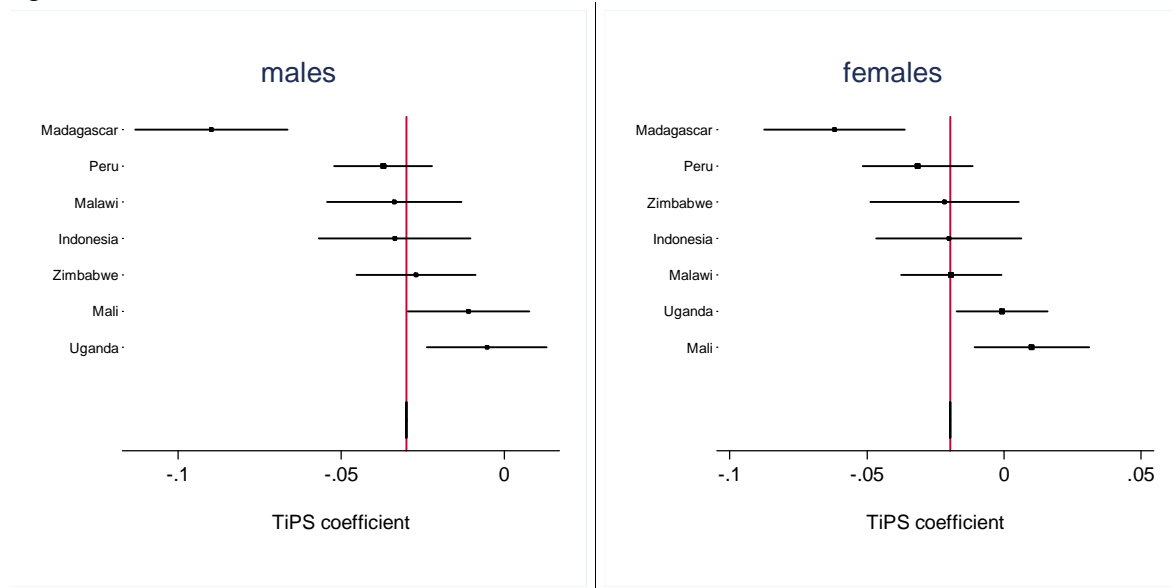


Figure 4

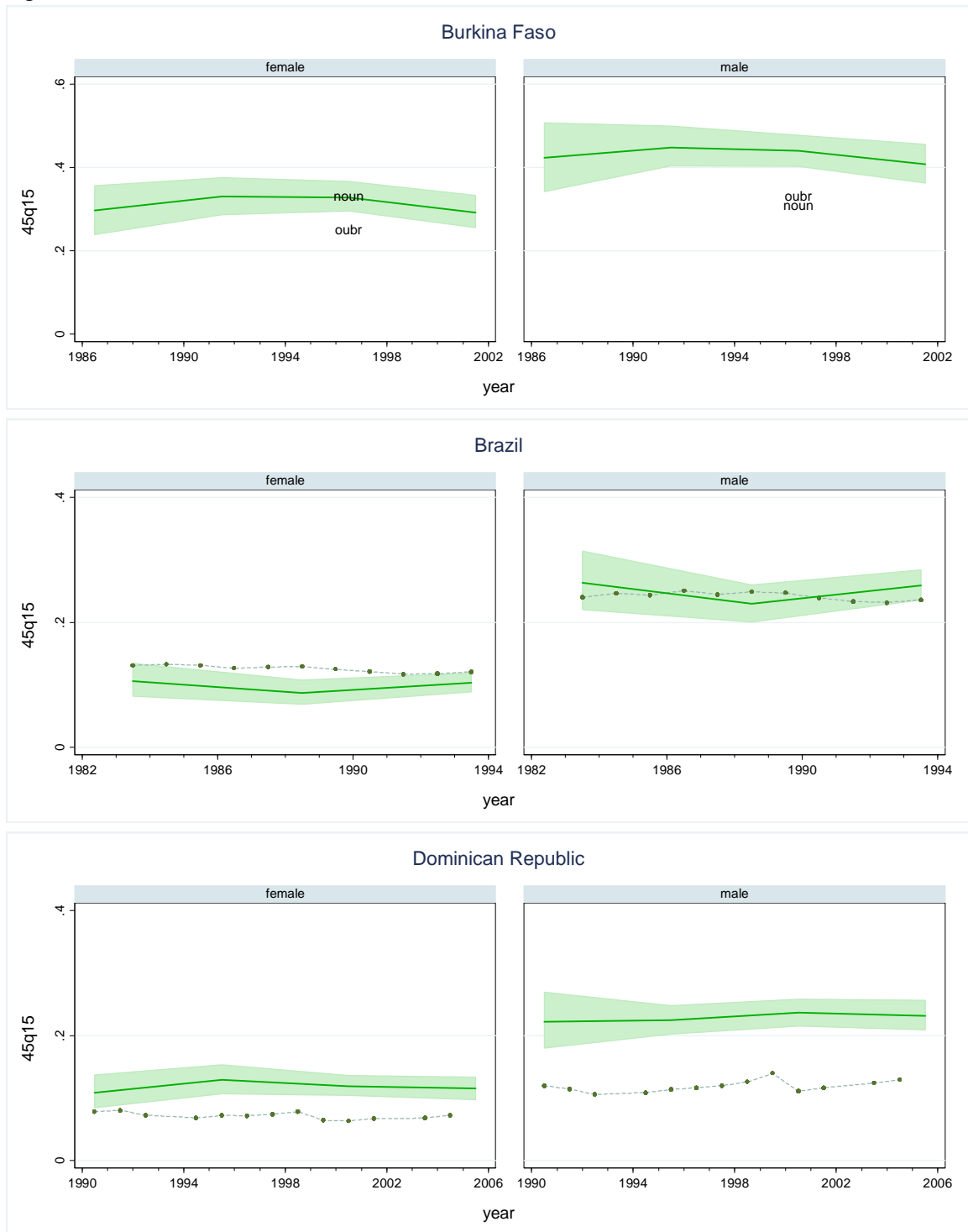


Figure 4, cont.

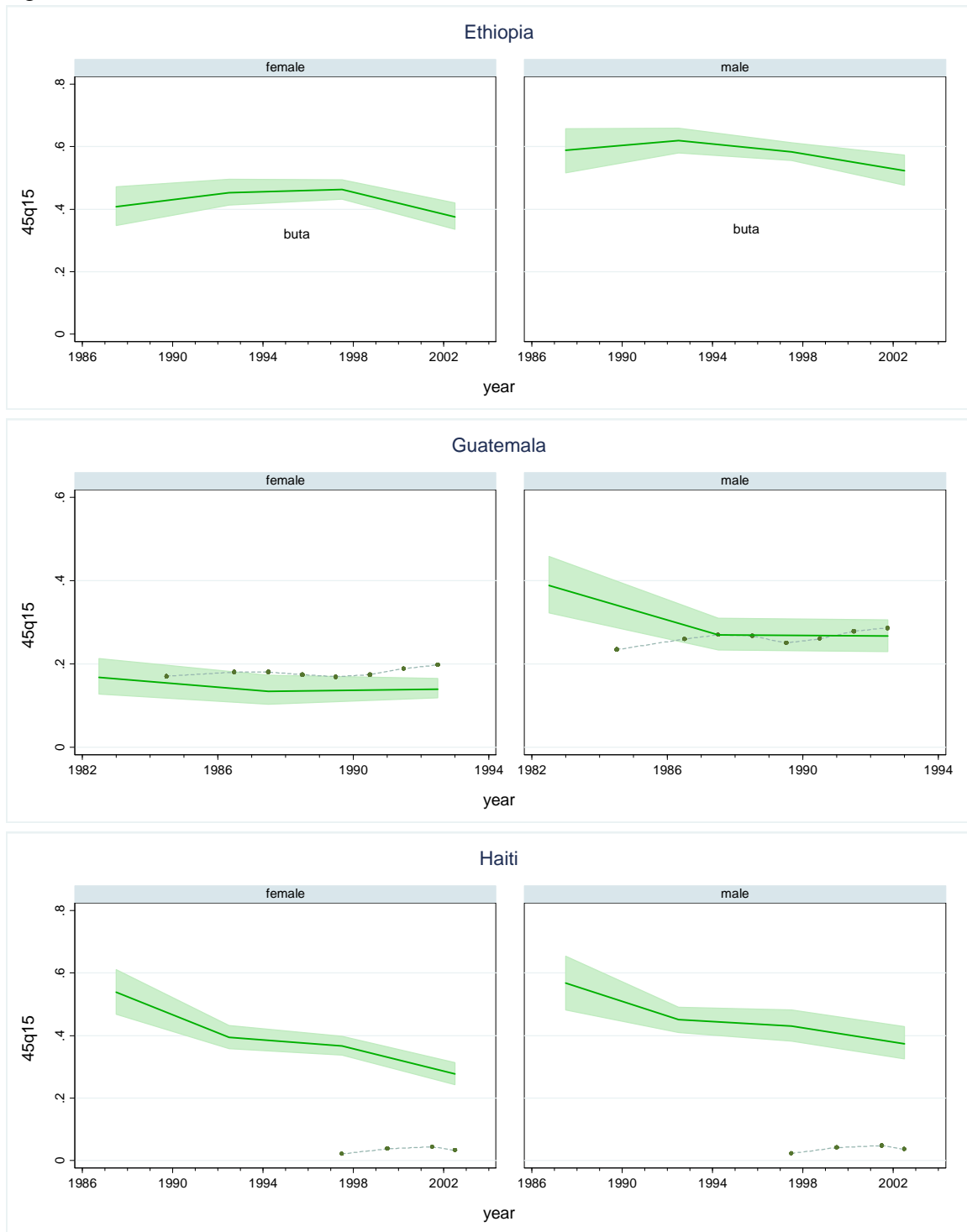


Figure 4, cont.

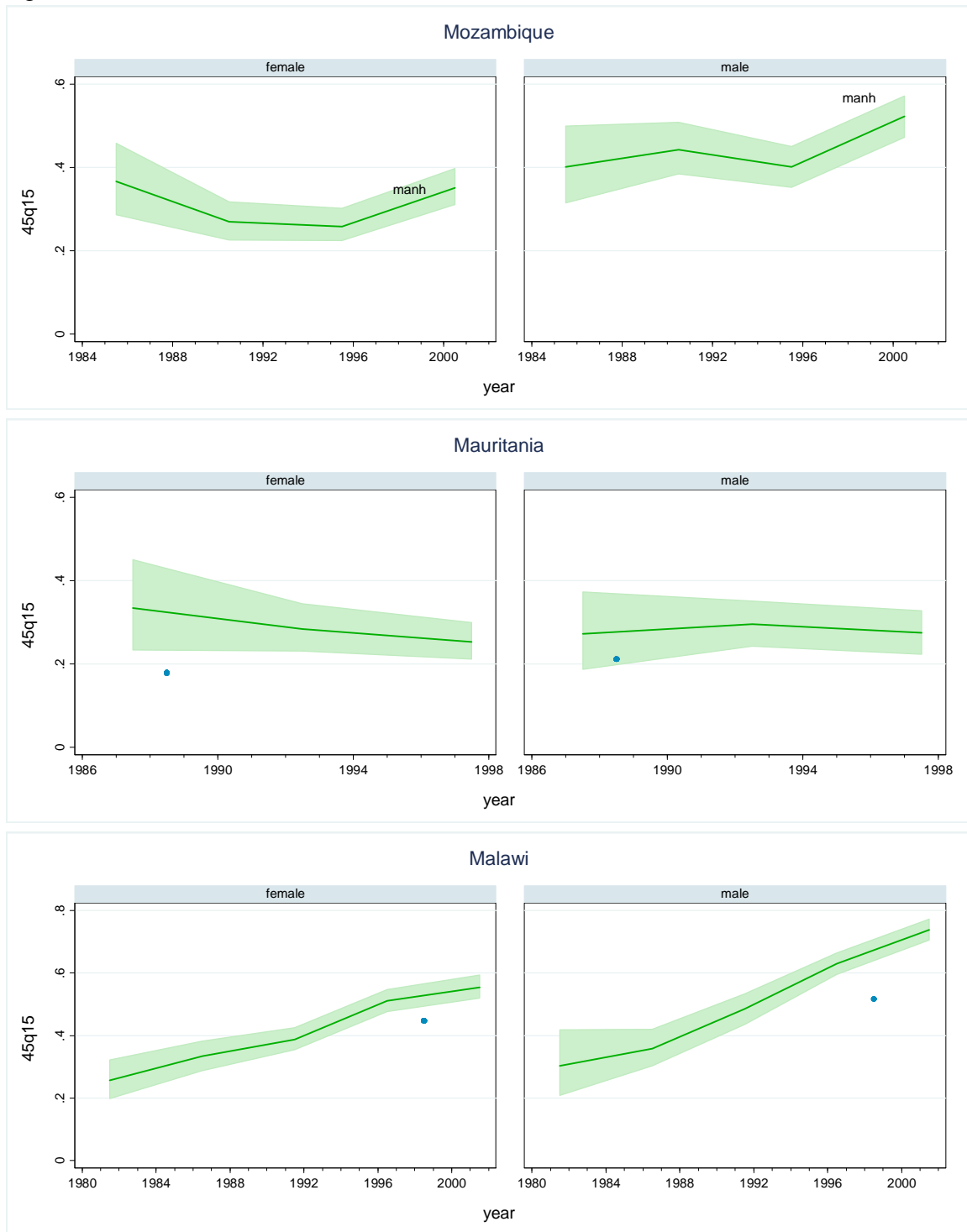


Figure 4, cont.

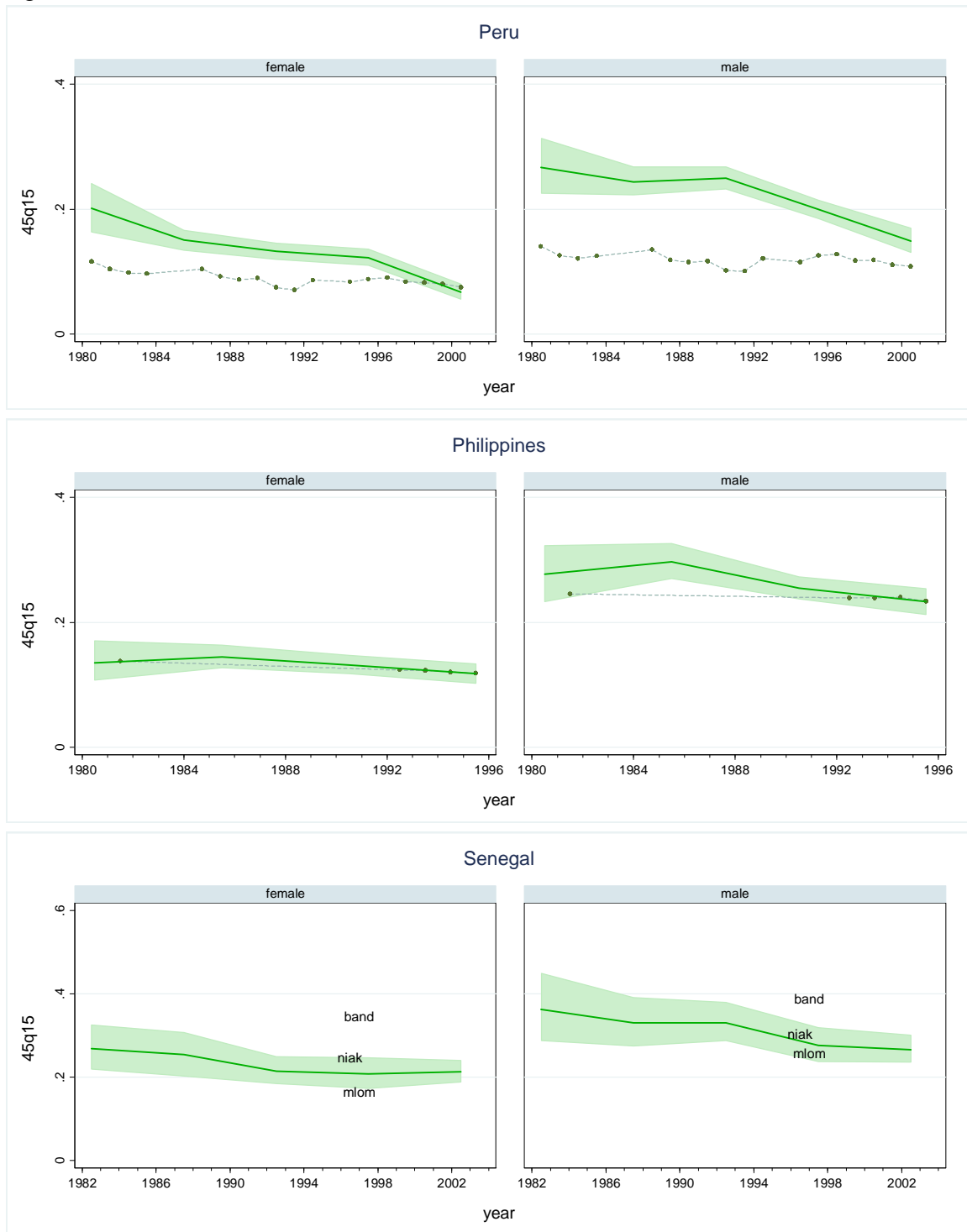


Figure 4, cont.

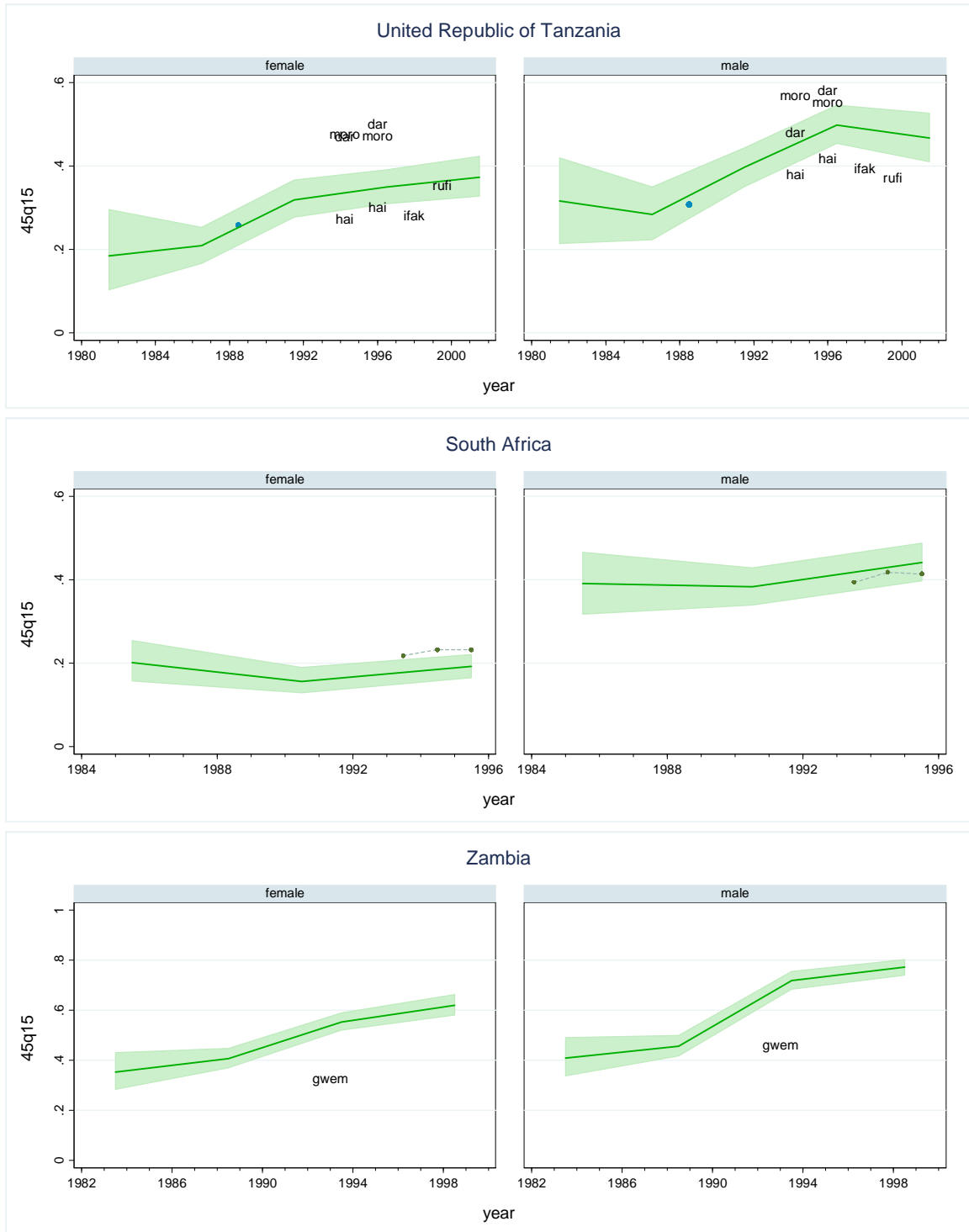
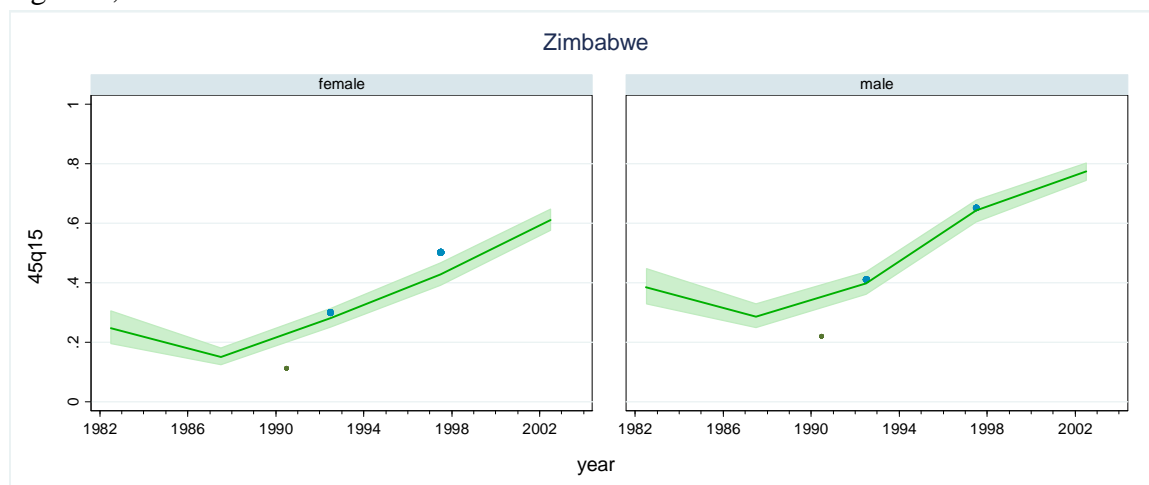


Figure 4, cont.



— CSS     ● Census  
- - - ● - - - VR     xxxx = DSS site

Notes: Key to DSS site abbreviations

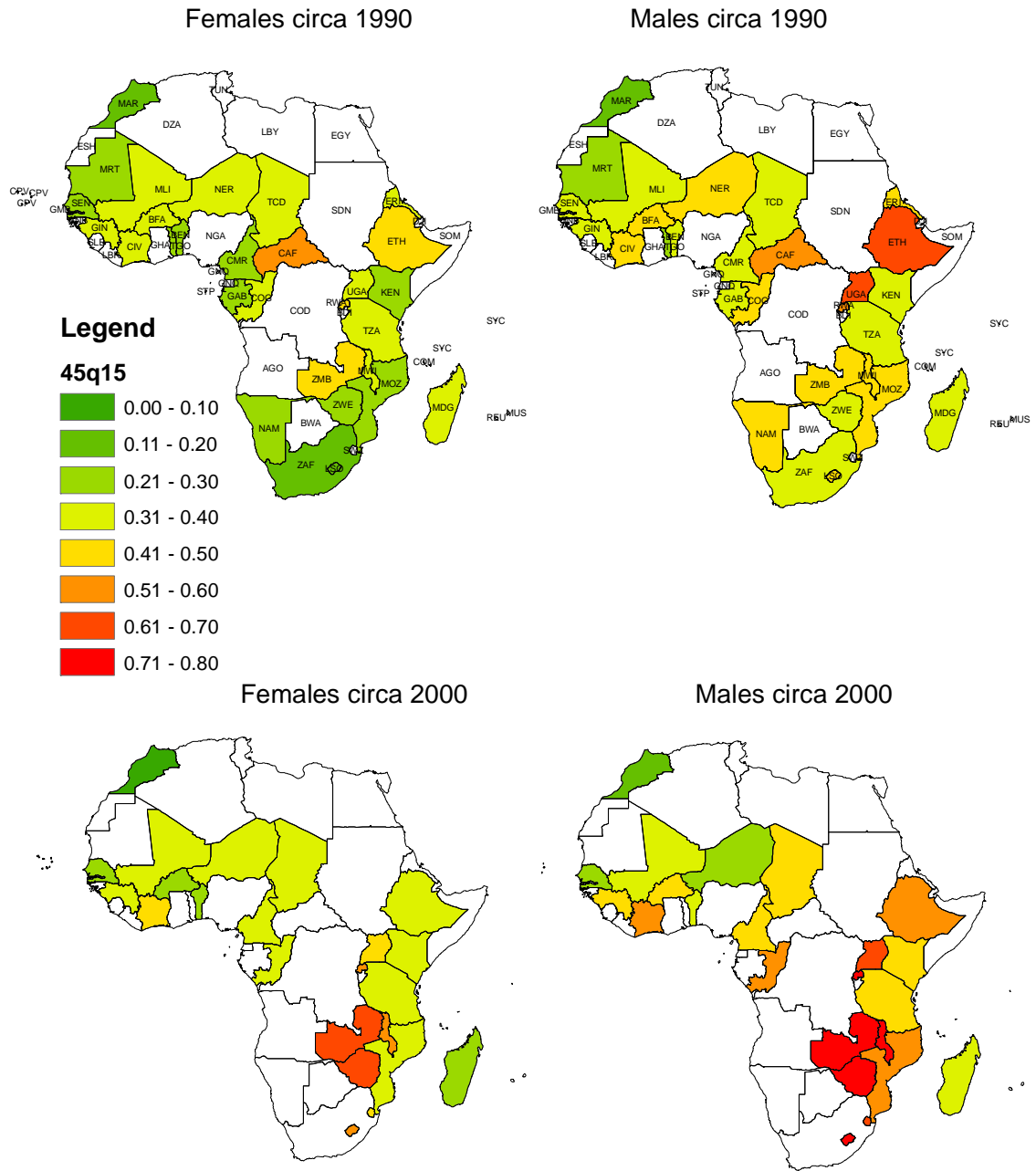
Site Name	Abbreviation	Site Name	Abbreviation
Agincourt	agin	Matlab-comparison	mat_c
Bandafassi	band	Matlab-treatment	mat_t
Bandim	band	Mlomp	mlom
Butajira	buta	Morogoro	moro
Dar es Salaam	dar	Navrongo	navr
Farafenni	fara	Niakhar	niak
Gwembe	gwem	Nouna	noun
Hai	hai	Oubritenga	oubr
Ifakara	ifak	Rufiji	rufi
Manhica	manh		

Data sources:

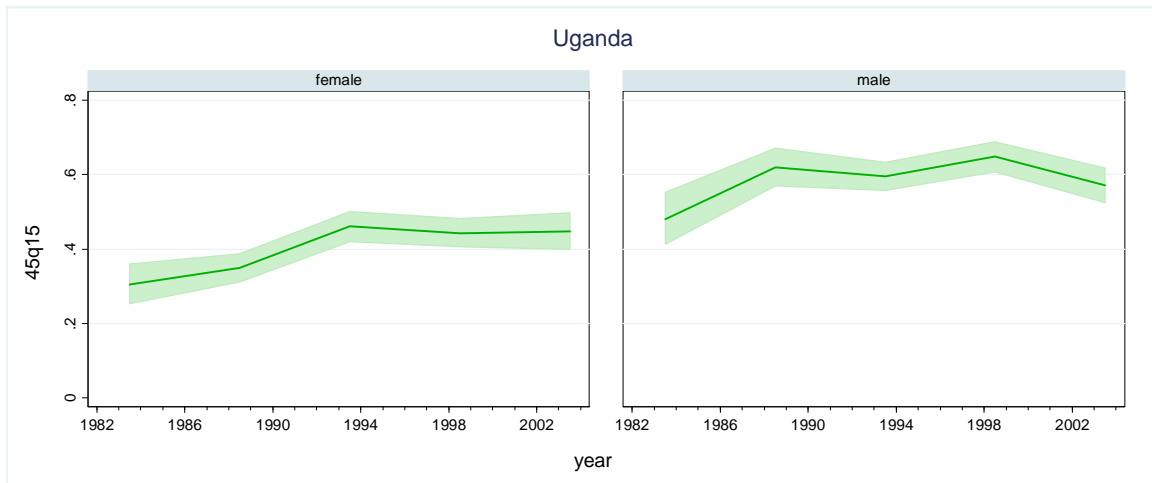
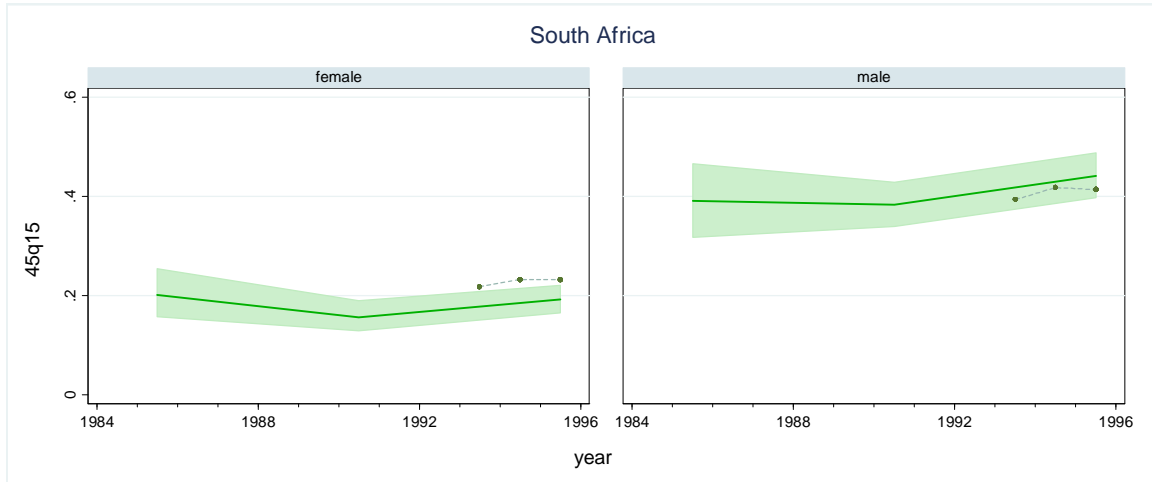
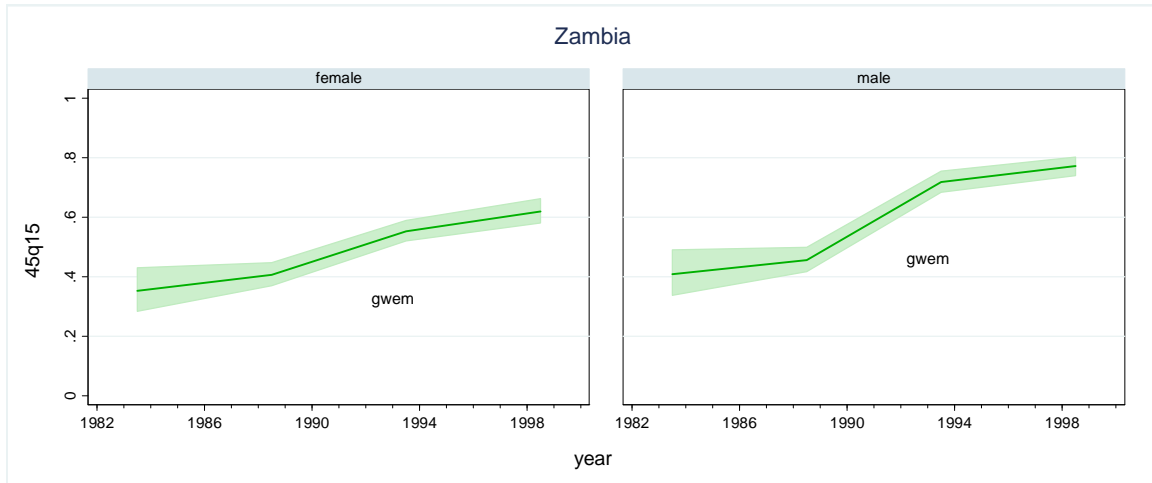
Vital Registration estimates: WHO Mortality Database 2008, UN Population Division World Population Prospects 2006  
 Demographic Surveillance Sites: INDEPTH 2002  
 Census estimates: WHO

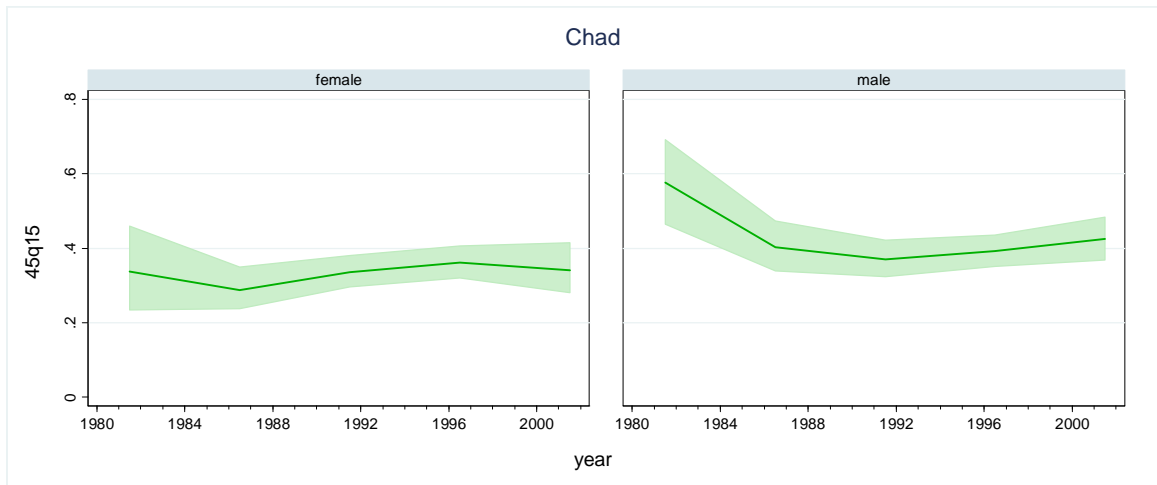
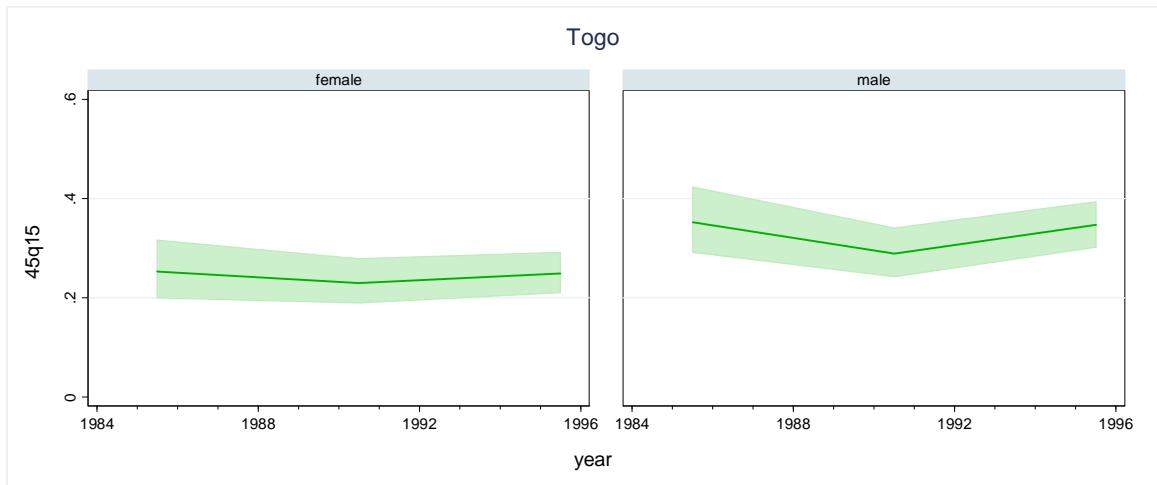
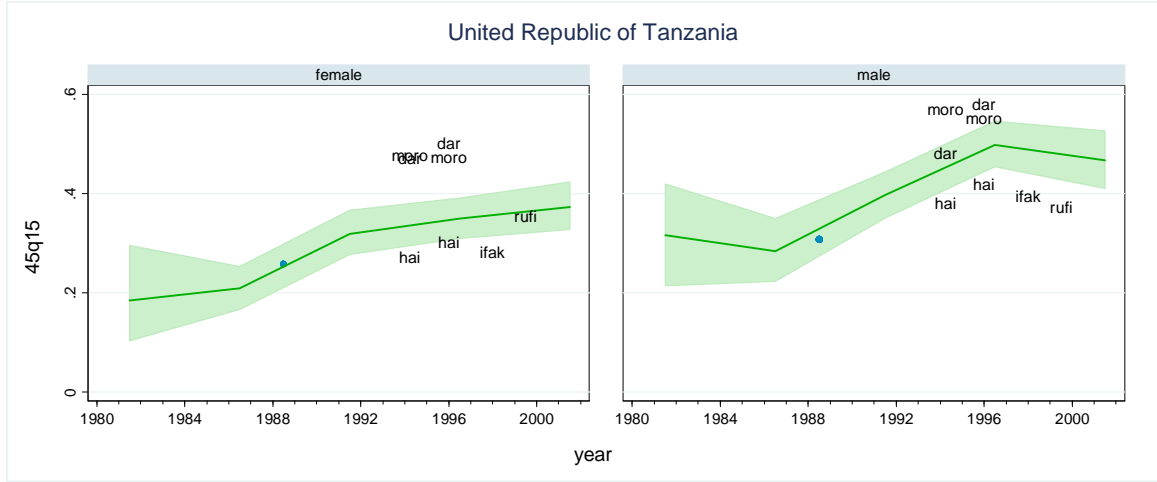


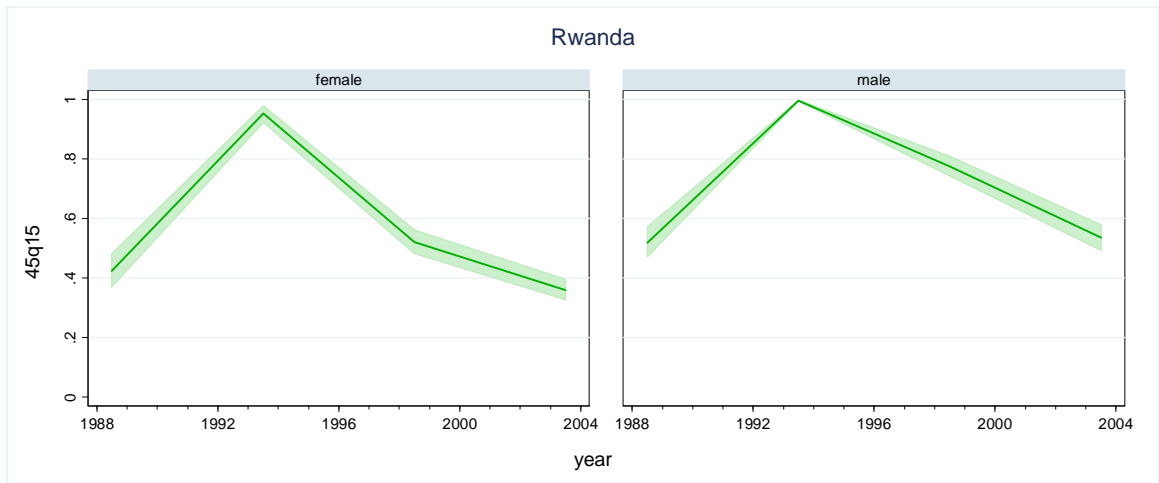
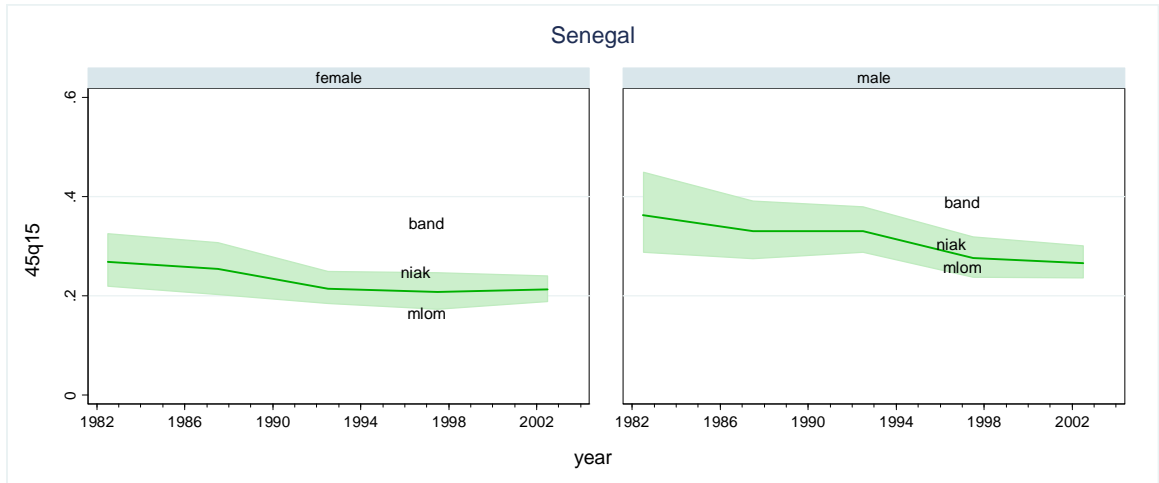
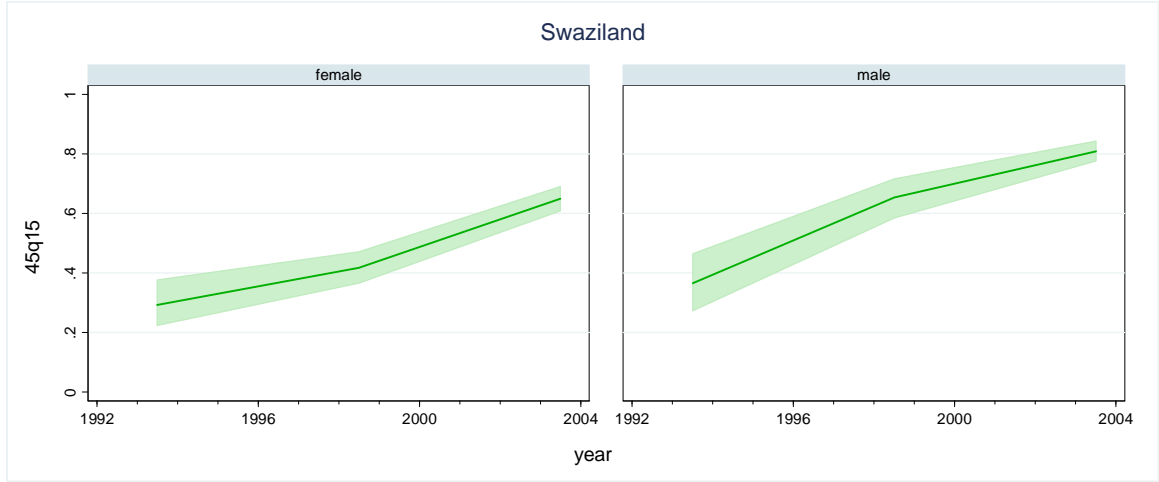
Figure 5



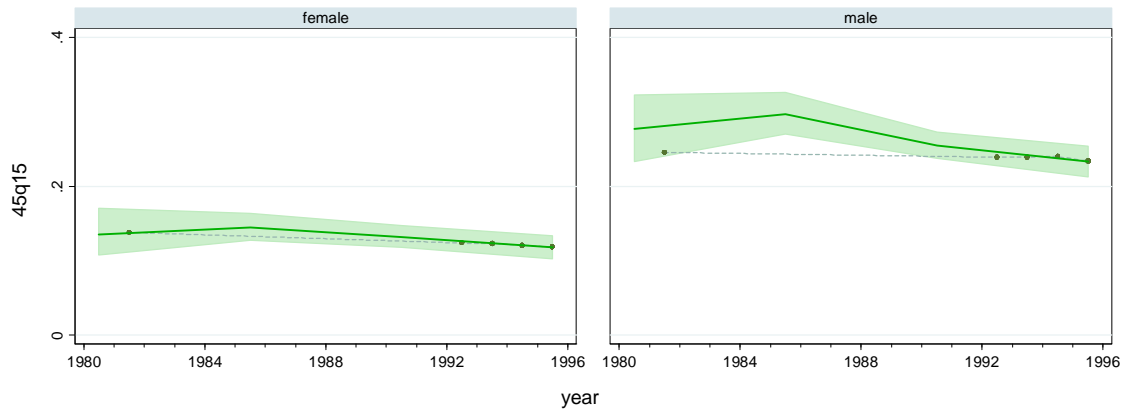
# Appendix



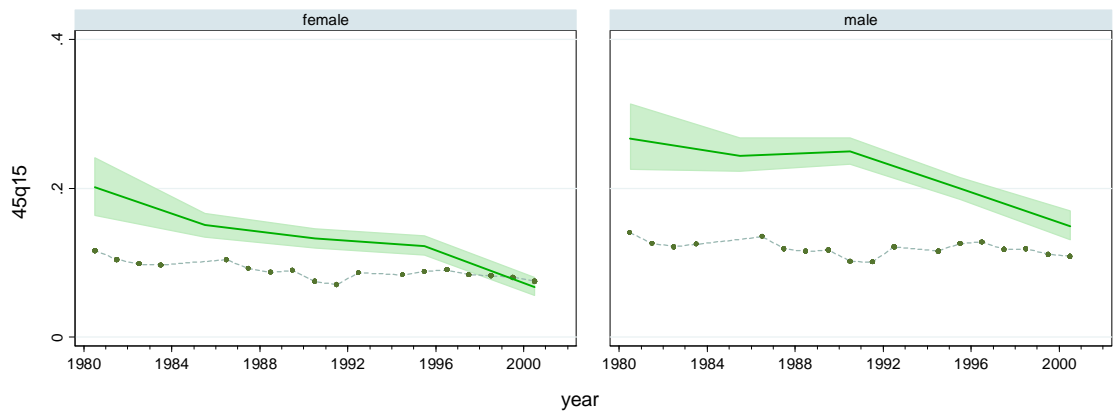




### Philippines



### Peru



### Niger

