TITLE: Reducing Ignorance about Adult Mortality: Improving Methods for Evaluating the Completeness of Death Registration

AUTHORS:
Christopher J.L. Murray, Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

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

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

Thomas Laakso, School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Alan D. Lopez, School of Population Health, The University of Queensland, Brisbane, QLD 4006, Australia

ABSTRACT:

Background:
One of the fundamental building blocks for determining the burden of disease in populations is to reliably measure the level and pattern of mortality by age and sex. Where well-functioning registration systems exist, this task is relatively straightforward. Results from many civil registration systems, however, remain uncertain due to a lack of confidence in the completeness of death registration. Death Distribution Methods (DDM) are a suite of demographic methods which attempt to estimate the completeness of death reporting. While widely applied and used, the methods have at least three types of limitations. First, a wide range of variants of these methods have been applied in practice with little scientific literature to guide their selection. Second, the methods have not been extensively validated in real population conditions where violations of the assumptions of the methods most certainly occur. Third, DDMs do not generate uncertainty intervals.

Methods and Findings:
In this paper, we systematically evaluate the performance of 312 variants of DDM methods in three different validation environments where we know or have strong beliefs about the true level of completeness of death registration. Using these datasets, we identify three variants of the DDMs which generally perform the best. We also find that even these improved methods yield uncertainty intervals of at least +/- 12%. Finally, we demonstrate the application of the optimal variants in eight countries.

Conclusions:
There continues to be a role for partial vital registration data in measuring adult mortality levels and trends, but such results should only be interpreted alongside all other data

sources on adult mortality and the face validity of the resulting levels, trends and age-patterns of adult death considered.

INTRODUCTION:

One of the fundamental building blocks for determining the burden of disease in populations is to reliably measure the level and pattern of mortality by age and sex. Simply knowing death rates at specific ages is of itself an important descriptor of the epidemiological situation in a population, given the strong age dependence of major diseases and injuries. After decades of effort and emphasis on improving survival among children, uncertainty about levels and trends in child mortality has been substantially reduced (although further improvements in knowledge are possible with better methods and wider access to survey and census data [1]. This is not the case with adult mortality, despite the focus on adult health outcomes in Millennium Development Goal (MDG) five (reducing maternal mortality) and MDG six (HIV, tuberculosis and malaria).

Given the importance of estimating underlying mortality rates in order to more reliably describe the burden of disease in populations, particularly for populations where the routine registration of deaths functions poorly, methods have been developed to more successfully exploit the substantial amount of information on the survival of siblings that has been collected in large scale global survey programs [2]. For many developing countries, the mainstay for adult mortality measurement including the maternal mortality rate remains civil registration systems. Over 50 developing countries annually report death statistics to the World Health Organization or the United Nations Statistical Office [3,4]. Results from using civil registration systems, however, remain uncertain due to a lack of confidence in the completeness of death registration and the accuracy of reports about age at death.

Beginning in the 1960s and 1970s, methods were developed by demographers in an attempt to estimate the completeness of death reporting, either in civil registration systems, or in censuses and surveys [5-10]. These methods, known in the literature as Death Distribution Methods (DDM) are effectively based on a comparison of the age distribution of recorded deaths with the age distribution of the population in which the deaths occurred. In order to satisfy basic demographic theory about forces of population growth, the methods are dependent on assumptions about the stability of populations, population growth and the extent of age misreporting. These methods have been widely applied to census and vital registration data in the literature and are used for nearly 100 countries by WHO to monitor adult mortality [11-16]. While widely applied and used, the methods have at least three types of limitations. First, a wide range of variants of these methods have been applied in practice with little scientific literature to guide selection of these variants. Second, the methods have not been extensively validated in real populations in the presence of measurement error. The only validation study [17] found large variation in results when applied to high-income countries where registration is thought to be complete. Third, DDM methods are grounded in mathematical not statistical relationships and thus do not generate uncertainty intervals for the estimated completeness of death registration.

In this paper, we systematically evaluate the performance of 312 variants of DDM methods. We use three different validation datasets where we know or have strong beliefs about the true level of completeness of death registration. Using these datasets, we identify improved DDM methods, characterize their uncertainty in different settings and then illustrate their applicability in developing countries.

METHODS

Three DDM Families

The three families of methods used for assessing the completeness of death registration are generalized growth balance (GGB), synthetic extinct generations (SEG) and a hybrid of the two approaches GGBSEG (Figure 1). Appendix A provides a brief summary of these methods and the mathematical relationships that underlie them. All that these methods require as input are age distributions of population from two censuses and the deaths registered between the censuses by age. The methods are normally applied separately to males and females. In some cases, instead of death captured by vital registration systems, deaths reported in a census in the last 12 months have been used. All three families of methods ultimately yield a correction factor which can be multiplied by the observed adult death rates to get the corrected adult death rates (see Appendix A). SEG methods yield an estimate of the completeness of death registration relative to the two censuses. GGB methods and the related GGBSEG yield an estimation of the completeness of census 2 relative to census 1 as well as the completeness of death registration relative to the censuses. In theory, GGB and GGBSEG should perform better in the presence of differential completeness of the two censuses.

Based on practical experience, results of applying DDMs to population and death data for all adult age-groups can give findings that lack face validity. Demographers often age-trim, namely drop the older and/or younger age-groups in the estimation of the correction factor for observed death rates. This practical approach has a sound theoretical basis: the effects of random fluctuations in the number of deaths or population in some age-groups, age misreporting and migration may vary substantially at older or younger age-groups. While age-trimming is widely practiced, there are no published studies which systematically evaluate the performance of different age-trims for the three families of DDMs. We have computed 78 age-trims for each of the three families. These 78 age-trims were chosen to cover all possible age-trims where at least five contiguous age-groups are used. We picked five as the minimum number of age groups required to give stable estimates for each of the methods. We identify each age-trim using the convention: *family a-k* where *family* is either GGB, SEG, or GGBSEG, *a* is the start of the age interval and *k* is the start of the last five year age-group included. In this article, we define "method" to mean the specific combination of family and age-trim, so effectively we are evaluating 3 families x 78 age-trims = 312 different DDM methods. Application of all methods has been in Stata [18].

Creating or Identifying Validation Environments

Choice of the optimal DDMs including age-trimming can only be undertaken in settings where the analyst has a reasonable knowledge of the true correction factor that needs to be applied to the observed death rate. The real challenge in this research area is creating or identifying existing validation environments. We use three different environments each with their own advantages.

i) Micro-simulation model of a population of 10 million followed for a period of 150 years exposed to different levels of age-specific mortality, fertility and migration. The advantage of the micro-simulation environments is that the analyst controls all aspects of population dynamics and measurement error and thus truth is known with certainty.

ii) US counties 1990-2000 provide a large set of populations with a large range in size, immigration and emigration rates where it is reasonable to assume that the relative completeness of vital registration relative to the 1990 and 2000 censuses is close to 100%.

iii) High-income economies as designated by the World Bank with populations greater than 5 million from 1950-2000. This group represents a much narrow range of migration rates, larger population sizes in countries with mature death registration systems.

Population and Measurement Micro-Simulation

Using micro-simulation to study the performance of DDMs requires two interconnected models: a population micro-simulation model and a measurement micro-simulation model. Figure 2a provides a schematic of the population model where individuals are exposed over time to age-specific risks of death, fertility and migration. Mortality and fertility rates were modeled based on trends in mortality and fertility in the US during the 20^{th} century. These were applied to an initial population age distribution from Sweden,

1751. The effect of fertility and mortality evolution over time on the population age-structure is illustrated in Figure 2b. After an initial period of approximately 75 years, the age distribution evens out and becomes smooth again. Using this population model, we created 11 different population scenarios based on different levels of mortality and fertility. For each mortality-fertility scenario, we added three scenarios of net immigration with rates of 1, 10 and 50 per thousand, three scenarios of net emigration with rates 1, 10 and 50 per thousand, and one scenario with no migration. The age-*pattern* of migration, however (illustrated in Figure 2c), is constant in each case of net migration and based on the average of a geographically diverse selection of countries with complete migration data as reported in the 1989 Demographic Yearbook [19]. In all, we generated 77 mortality-fertility-migration population scenarios with data on roughly 10-15 million individuals in each. Various demographic characteristics of each population scenario are shown in Table 1.

For each population scenario, we applied a measurement micro-simulation model where census 1 is taken at time *t*, registration of deaths occurs from time *t* to *t*+10 and census 2 is taken at time *t*+10. Individuals have probabilities of being included in the two censuses of *c1* and *c2* and, if they die, of being registered of *v1*. Further, each individual's age in each measurement is recorded subject to two types of age-misreporting: stochastic and systematic. Stochastic age-misreporting is captured as a random draw for each individual for each measurement from a normal distribution with mean zero and variance σ_1^2. Systematic age-misreporting is captured by the function: $a_m = a_t + a_t\beta$ where a_m is the misreported age, a_t is the true age, and β is drawn from a normal distribution with a mean μ and variance σ_2^2. We vary the choice of $c1$, $c2$, $v1$, and the nine parameters defining the age-misreporting distributions $\sigma_1^2(c1)$, $\sigma_1^2(c2)$, $\sigma_1^2(vr)$, $\mu(c1)$, $\mu(c2)$, $\mu(vr)$, $\sigma_2^2(c1)$, $\sigma_2^2(c2)$, and $\sigma_2^2(vr)$, randomly generating 2000 different measurement scenarios for each of the 77 population scenarios. Table 2 summarizes the ranges for the nine parameters governing the measurement process that we have used in the simulations. Because we believe that age-misreporting is likely to be culturally determined, we have built strong correlations into the selection of age-misreporting variables between any given measurement micro-simulations (i.e. age misreporting in one measurement event such as census 1 is similar to age misreporting in another measurement event in the same population). The choice of the ranges sampled in Table 2 is based on our review of the literature [20,21]. In total, we have generated 154,000 sets of two censuses and death registration data over 10 years where we know the true death rate and the observed death rate and thus the correction factor that DDMs should generate.

US Counties 1990 to 2000

Our second validation environment is US counties 1990 to 2000 where census age counts and death registration are available. We use the 2,072 counties or merged county units developed to assure a minimum population size in each aggregate [22-24]. Table 3 summarizes the range of population sizes, mortality, fertility and migration observed across US counties. We assume that due to rigorous enforcement of census and registration laws that relative completeness of death registration is effectively 100% in all counties. To remove the effect of small numbers on these methods, we also present results for the 534 counties with a population greater than 100,000. Immigration and emigration data are based data from the United States Internal Revenue Service 2007 [25], which tabulates the number of exemptions (an estimate of the number of individuals) that move from each county to every other county by matching the Taxpayer Identification Number and comparing zip codes of filing addresses from one year to the next.

High-Income, Large Countries 1950-2000

Our third validation environment, following the work of Thomas & Hill (2007) [17], is high-income countries with populations greater than 5 million from 1950-2000. We have identified 195 periods[1] across 17 countries[2] where census data are available at the beginning and end of the period and death registration data is available for part or all of the intermediate years. The data sources include the United Nations Demographic Yearbook [26] and the WHO mortality database [27]. We have applied the best performing method for each family of DDMs to all 195 combinations of censuses and death registration within the matching intercensal period, yielding 195 estimates of relative completeness for each DDM for these countries. For each of them, we assume as in the US counties that relative completeness of death registration is very close to 100% since social and legal structures in place for several decades mean that it is extremely difficult to dispose of a corpse without legal registration of death.

---

[1] Periods were defined by pairing a census with each of the two subsequent censuses in time; this yielded two intercensal periods (for example, a census in 1970 would be paired with the 1980 census to create one period and the 1990 census to create a second period, the same would be done with 1980-1990 and 1980-2000 and so on). As required to apply the methods, deaths during the intercensal period were averaged to create average annual deaths within each period.

[2] The 17 countries include: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Israel, Italy, Japan, Republic of Korea, Netherlands, Portugal, Saudi Arabia, Spain, Sweden, Switzerland, Great Britain, and United States of America.

Evaluation the Performance of Different DDMs

Each method yields a correction factor which can be multiplied by the observed death rate to get a corrected death rate. Appendix A provides a formula for each of these correction factors by DDM family. For convenience and interpretation, we define 'relative completeness of death registration' to be the inverse of the correction factor from each family. It is important to note that the assumptions in each family that are incorporated into the correction factor are different. Nevertheless, we ultimately want DDMs that yield the correction factor closest to the true value needed to correct the observed death rate for a population to equal the true death rate. For ease of communication, we prefer to evaluate DDMs using the inverse of the correction factor, or relative completeness of death registration (RC). For each validation environment, we compare estimated RC to the true or assumed RC. The difference between RC(estimated) and RC(true) is the error in the estimated relative completeness. We use average relative error as a metric of performance of a given method applied in a given validation environment. More formally:

$$\textit{average relative error}_{tme} = \frac{\sum_{i=1}^{N} \frac{\left|(RC(estimated)_{tmei} - RC(true)_{ei}\right|}{RC(true)_{ei}}}{N_e} \qquad (1)$$

Where *tm* represents the age-trim *t* used from family *m* (GGB, SEG or GGBSEG), *i* indexes each simulated population, county or country from validation environment *e* and *N* is the total number of such populations in that environment. We choose optimal DDMs for each family of methods by minimizing the average relative error in the three validation datasets.

Application to Selected Developing Countries

As with the high income countries, in applying the methods to developing countries we create periods by pairing each census with the two subsequent censuses. We then apply the optimal age-trims[3] for each of the three families of methods to the resulting census pairs and the intercensal average annual deaths from death registration or census/survey data on household deaths[4] found in the UN Demographic Yearbooks [26], IPUMS [28], and WHO [27] mortality databases for 1950-2000. In total, this yielded roughly 1,000 estimates of each optimal DDM. For illustrative purposes, we present our results in detail for 6 developing countries and contrast them with results from 2 high income countries.

RESULTS

The performance of all possible age-trims for the three families of DDMs in the three validation datasets is summarized in Table 4 (which lists the top 5 and worst five age-trims for each method). The full results for every age-trim can be found in Appendix Table 1. The results in all validation datasets demonstrate high variation in performance across different age-trims. This variation ranges from 4.8% to 142% in terms of average

[3] In some cases, the age groups in the data did not allow for the use of the optimal age-trim. For example, if the open interval for deaths was 70+, we could not apply any age trims that went above the 65-69 year age group. In these cases, we applied the best performing age-trims possible given the age groups present in the dataset.

[4] When the death data available was from censuses or surveys which asked about household deaths in the last 12 months, we computed average annual deaths by calculating the death rates at the time of the second census (or survey), and applied them to the average person-years lived assuming geometric population growth in the intercensal period.

relative error. Clearly the key determinants of the performance of each family of DDMs vary profoundly according to which age groups are included in the estimation process. There is much greater variation across age-trims than there is across families of DDMs. As Table 4 shows, we have computed the average relative error for each age-trim in each validation environment and ranked the trims within each environment. The minimum average rank across the three environments yields the best performing method. For SEG the optimal age-trim is SEG 60-80; this was the best in the simulated populations, U.S. counties and high income countries. For GGB, the results across validation datasets appear to be more mixed, but GGB 30-80 performed best on average. Finally, GGBSEG 50-70 performed best on average across the three validation environments.

Given the closest test to national application is the high-income countries, all three optimal versions of the three DDM families perform relatively well with similar average relative error in this setting. GGB 30-80 does slightly better than the optimal age-trims in the other families. Of note, the previously reported sensitivity of SEG to migration in high-income countries [17] appears to be largely attenuated in SEG 60-80. We focus on the three optimal methods, SEG 60-80, GGB 30-80 and GGBSEG 50-70 for the rest of the analysis.

The simulation dataset provides an opportunity to investigate how error and estimated relative completeness are associated with factors such as the levels and trends in mortality, fertility, migration and age-misreporting. Table 5 shows regression results of error in relative completeness for each of the three optimal methods regressed on the 6 age-misreporting variables (stochastic and systematic age-misreporting variables for each of the two censuses and the vital registration system) and migration rate with and without fixed effects for the 77 population scenarios. Overall, the regression results show large effects in all three optimal trims for age-misreporting. Stochastic age-misreporting has an important effect but the effect of systematic age-misreporting is much larger, judging by the t-statistics. Of particular importance are differences in the systematic age-misreporting variables across the two censuses and vital registration (separate regressions not shown). The greater the difference in age misreporting across the censuses and VR, the greater the error. Adding population fixed effects increases the R-squared of the regression, indicating that the parameters defining the population (mortality, fertility and migration) along with age-misreporting together explain a total of 78.1%, 94.4% and 90.5% percent of the variation in the error of the SEG 60-80, GGB 30-80 and GGBSEG 50-70 methods, respectively. With fixed effects, migration is not significant as expected since the 77 fixed effects capture the unique effects of the combination of mortality, fertility and migration. The coefficient on migration in the model without fixed effects is relatively small in comparison to the error generated by the age-misreporting parameters. In part, this result may be due to the selection of the optimal age-trims which tend to minimize the impact of migration already.

We also analyzed the performance of the age-trims in the simulations according to level of migration. Though the best performing age-trims as measured by average relative error are consistent across varying levels of migration for SEG and GGBSEG, they are not so for GGB. For SEG and GGBSEG, the age-trim 60-80 has the smallest average relative error in all migration scenarios. For GGB, in the absence of net migration, the age-trim 5-65 has the smallest average relative error, while the age-trims 45-65, 50-70 and 40-65 are the better performers under positive or negative net migration scenarios.

Uncertainty in the estimated relative completeness is large. Figure 3 shows in the three validation datasets the relationship between the estimated relative completeness and the true relative completeness for the optimal method in each of the three families. The variation in estimated relative completeness increases as true coverage approaches 100%. When the error in relative completeness is expressed in relative terms by dividing by the true level of relative completeness, error remains constant as a function of true completeness (not shown). The uncertainty in the simulated datasets and the US counties is similar but it is smaller in large high-income populations which may reflect larger numbers, lower migration rates and smaller age-misreporting at the national level.

Figure 4 shows 8 examples of the application of the three optimal methods to select countries over time. In each panel, for each pair of censuses and vital registration data, we show the results of the three methods in terms of relative completeness. For comparison we have also compared registered deaths 0-4 to estimates of 0-4 deaths based on systematic review of all data sources [1]. We include graphs from two high income countries, Canada and Switzerland, to show that these methods can be consistent and, as we assume death registration is complete in these countries, accurate. The graph for Mexico suggests that vital registration, at least for adults, has been relatively complete since 1970, but there is clearly more noise in these estimates than for the high-income countries shown. The Philippines shows similarly noisy estimates, with both improvement and decline suggested. Knowing the uncertainty inherent in these methods, however, it is unclear that these are true trends. In Thailand, registration for adults is estimated to be between 80 and 90 percent complete during the 1980s and 1990s, and the methods are fairly congruous. Paraguay is another example where registration completeness has been relatively constant over time and where the methods seem to be fairly consistent with one another. Tunisia illustrates an example where death registration has clearly improved over time, from nearly 50% in 1960 to complete by 1980. Finally, the graph for Korea reminds us that these methods are applicable not only to data from vital registration systems, but also can be applied to survey sources of death data. Across all the countries, it is not clear that any one family of DDMs is best or most consistent, although SEG appears to be slightly more stable.

DISCUSSION

This first systematic analysis of the three families of DDMs for the estimation of the completeness of death registration demonstrates that the choice of age-trimming has a profound effect on the performance of these methods. Based on three different validation datasets, we believe SEG 60-80, GGB 30-80, and GGBSEG 50-70 are the best methods that can be currently used to estimate relative completeness of death registration. We recommend that analysts apply all three methods and look at the consistency of results. The combination of the three optimal DDMs will yield much better results than current practice of application of DDMs without optimal age-trimming. Selection of optimal age-trims has also substantially reduced the bias associated with migration reported in previous work.

Published studies and national statistical reports apply these methods and provide results without uncertainty bounds. In the three different validation environments, estimated relative completeness for the best variants has a minimum average relative error of 4.8%, and this can be as high as 11.3% depending on the method used. It appears that the underlying stochastic processes in censuses and death registration, including age-misreporting, have led to a component of uncertainty that cannot be eliminated. While usage of partial death registration is useful for estimating mortality levels among adults, the application of DDMs, even from the optimal age trims we have suggested here, should be interpreted with considerable caution; relative completeness of registration is likely to be at least +/- 12 around the estimated level, and perhaps considerably more. This level of uncertainty is likely to mean that while DDM correction methods could be useful in estimating levels, they are unlikely to be as useful for estimating mortality change. For example, Lopez and colleagues [29] estimate that *45q15*[5] for females in Paraguay declined by 8 per 1,000 over the period of 1990-2001. According to our analysis, a +/-12 point uncertainty interval around the SEG 60-80 estimated relative completeness of 74% for Paraguay in the late 1990s would yield an uncertainty interval around predicted *45q15* for females of between 93 and 127 per 1,000, a spread of 34 points per 1,000. Detecting the decline in adult mortality that is estimated to have

[5] 45q15 is life table notation indicating the probability that a 15 year old would die by age 60 if mortality rates of the reference period (in this example, the late 1990s) were constant. It is a commonly used summary measure of adult mortality.

happened in Paraguay during this time period would not be possible given the uncertainty inherent in the DDMs.

Our working hypotheses in applied work have been that (1) the completeness of adult death registration is always greater than or equal to the completeness of child death registration given the greater ease of disposing of infant or child remains without notice of legal authorities compared to adult remains. In addition, we have operated under the assumption that (2) the evolution of social and public institutions leads to stronger civil registration that will improve both adult and child death registration and thus generate a high correlation between adult and child completeness. Application of our optimal DDMs, however, provides indications that assumption (2) may not be entirely accurate. There are a number of developing countries in Latin America and South-East Asia where adult registration appears to be complete, but child registration varies from complete to less than 50%. There may well be considerable variation across countries in the time lag between achieving complete or near complete adult death registration and the same for children.

Given the residual uncertainty in optimal DDMs, there may be a bigger role for direct measurement of relative completeness through surveys or censuses. Two methods deserve broader consideration. First, some surveys such as Thailand's 1995-1996 Survey of Population Change [30] have asked households about deaths in the last 12 months and whether the death was registered. It is possible that in countries where death registration is legally required, the reported levels of death registration may be inflated. Nevertheless, this avenue of measurement could be further refined. Household respondents could be asked if deaths also occurred in hospital, for example. The number of hospital deaths recorded by the health information system could be examined to cross-validate household responses. A second strategy would be to apply capture-recapture or dual-record methods [31,32] to civil registration deaths and deaths reported by households in a time period prior to a survey or census. Capture-recapture methods require matching of individual deaths so this effort can be time consuming. Direct measurements of completeness using this approach have been used in the Chinese Disease Surveillance Point System [33] and at Demographic Surveillance Sites in Kenya [34], as well as with recent work in Thailand with the most recent 2006 Survey of Population Change [35]. More experience with both types of approaches may strengthen our capacity to track the completeness of death registration.

The analysis in simulated populations of the profound impact of stochastic and systematic age-misreporting has a more general implication. Preston and others [36] have pointed out that even in complete death registration systems, age-misreporting can bias the measurement of death rates by age. In a typical developing country with a young age-structure, even stochastic age-misreporting will lead to over-estimation of death rates at younger ages and under-estimation of death rates at older ages. The sensitivity of completeness estimates from DDMs to age-misreporting compounds this problem. There will need to be renewed efforts to measure the extent of stochastic and systematic age-misreporting and provide tools for correcting the bias in observed death rates. This bias is likely present in all available national life tables at present.

Given the increasing availability of other measurements of adult mortality such as corrected sibling survival, corrected death registration data should be interpreted in the context of all other data sources. In the arena of child mortality, it is now standard practice [1] to examine all data sources for a country over time and generate a composite estimate of levels of and trends in child mortality. We believe with improved DDMs, there continues to be a role for partial vital registration data in measuring adult mortality levels and trends. But such results should only be interpreted alongside all other data sources on adult mortality and the face validity of the resulting levels, trends and age-patterns of adult death considered.

ACKNOWLEDGMENTS

The authors wish to thank Ken Hill for helpful comments on the research, Sandeep Kulkarni for assistance with the US County dataset, and Stanislava Nikolova for assistance with referencing.

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: CJL Murray, JK Rajaratnam, J Marcus and AD Lopez contributed to the writing of the manuscript. All authors provided input into the analysis. Christopher JL Murray and Alan D Lopez conceived of the research idea.

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

Appendix A

Generalized Growth Balance

The Generalized Growth Balance (GGB) method extends the Growth Balance Equation developed by Brass in 1975. The original method employs the demographic relationship

$$\frac{N(a)}{N(a+)} = r + \frac{D(a+)}{N(a+)}$$

where $N(a)$ is the number of people turning age a in the interval under observation, $N(a+)$ is the total population above age a, r is the stable population growth rate, and $D(a+)$ is the total deaths at age a and over. Generalizing the equation to non-stable populations, i.e. populations with non-constant growth, Hill estimated an r for each age group, yielding

$$\frac{N(a)}{N(a+)} = r(a+) + \frac{D(a+)}{N(a+)}$$

The method requires population by age at two points in time and the average annual deaths in between those points. If we have incomplete censuses and incomplete death registration, then we cannot observe $N1(a)$, the true population of people aged a at the first census, $N2(a)$, the true population of peopled aged a at the second census, or $D(a)$, the true number of deaths of people aged a in between the first and second censuses. We instead observe $N1^O(a)$, $N2^O(a)$, and $D^O(a)$, where

$$N1^O(a) = c_1 * N1$$

$$N2^O(a) = c_2 * N2$$

$$D^O(a) = v_1 * D(a)$$

In the above equations, c_1 is the completeness of the first census, c_2 is the completeness of the second census and v_1 is the completeness of death registration. Substituting these expressions into the Generalized Growth Balance equation allows us to estimate the relative completeness of the death registration to the censuses. After substitution and simplification, the equation becomes

$$\frac{N^O(a)}{N^O(a+)} = \frac{1}{t}\log\frac{c_1 N2^O(a+)}{c_2 N1^O(a+)} + \frac{\sqrt{c_1 c_2}}{v_1}\frac{D^O(a+)}{N^O(a+)}$$

By pulling the ratio $\frac{c_1}{c_2}$ out of the growth rate term and rearranging, we obtain

$$\frac{N^O(a)}{N^O(a+)} - r^O(a+) = \frac{1}{t}\log\frac{c_1}{c_2} + \frac{\sqrt{c_1 c_2}}{v_1}\frac{D^O(a+)}{N^O(a+)}$$

which defines a line with intercept $\frac{1}{t}\log\frac{c_1}{c_2}$ and slope $\frac{\sqrt{c_1 c_2}}{v_1}$. Least squares regression can estimate both the slope and intercept of this line and thus the relative completeness of the censuses and a correction factor for the mortality rates.

Synthetic Extinct Generations

The method of extinct generations, a precursor to the Synthetic Extinct Generations method (SEG), estimates the number of people aged a at time t, $N_t(a)$ by employing the following demographic relationship between $N_t(a)$ and $D_t(a)$, the number of deaths of people aged a at time t.

$$N_t(a) = D_t(a) + D_{t+1}(a+1) + D_{t+2}(a+2) + \cdots = \int_a^{\infty} D_{t+x-a}(x)dx$$

We can arrive at $N_t(a)$ either directly or indirectly by counting each person contributing to $N_t(a)$ when he or she dies [37]. The method of extinct generations is impractical for those who need the resulting death information now. Instead of waiting for cohorts to go extinct, SEG estimates the number of future deaths in the cohort by adjusting D_t at all ages above a by a growth factor. The new approach yields the equation

$$N_t(a) = \int_a^{\infty} D_t(x)e^{\int_a^x r(u)du}dx,$$

where the right hand side of the equation represents the estimated deaths at older ages to the current cohort $N_t(a)$ by using current deaths at each age $D_t(x)$ from death registration systems and age specific growth rates $r(u)$ that can be obtained from comparing cohort sizes across two censuses. Dividing the estimated $N_t(a)$ calculated using the right hand side of the equation by the $N_t(a)$ observed directly from the census gives the relative completeness ratio $\frac{v_1}{c_1}$.

Generalized Growth Balance-Synthetic Extinct Generations

If the two censuses used to estimate the age-specific growth rates in SEG have different levels of completeness, then SEG's final estimate of the correction factor will be biased. The combined GGB-SEG method attempts to correct this bias. Hill and Choi proposed multiplying the population numbers from the first census by the $\frac{c_2}{c_1}$ derived from the intercept of the Generalized Growth Balance equation. After this adjustment, the two censuses are complete with respect to each other and SEG will produce a less biased estimate of the correction factor.

References

1. 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.

2. Obermeyer Z, Park C, Rajaratnam JK, Gakidou E, Lopez AD, Murray CJL (2009). Measuring adult mortality using sibling survival: a new analytical method and new results for 38 countries, 1970-2005. PLoS Medicine (submitted).

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. 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.

5. Bennett NG, Horiuchi S (1981). Estimating the completeness of death registration in a closed population. Population Index 47(2): 207-221.

6. Brass W, Coale AJ (1968). Methods of analysis and estimation. In: Brass W, Coale AJ, Demeny P, Heisel DF, Lorimer F, Romaniuk A et al. The demography of tropical Africa. Princeton, NJ: Princeton University Press.

7. Hill K, Choi Y (2004). Death distribution methods for estimating adult mortality: sensitivity analysis with simulated data errors. Paper prepared for Adult Mortality in Developing Countries Workshop. Marin County, CA: The Marconi Center, July 8-11.

8. 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. (accessed January 7, 2009).

9. Preston S, Coale AJ, Trussell J, Weinstein M (1980). Estimating the completeness of reporting of adult deaths in populations that are approximately stable. Population Index 46(2): 179-202.

10. Preston SH, Hill K (1980). Estimating the completeness of death registration. Population Studies 2: 349-366.

11. Banister J, Hill K (2004). Mortality in China 1964-2000. Population Studies (Camb) 58(1): 55-75.

12. Elo IT, Preston SH (1994). Estimating African-American mortality from inaccurate data. Demography 31(3): 427-458.

13. Feeney G (2001). The impact of HIV/AIDS on adult mortality in Zimbabwe. Population and Development Review 27(4): 771-780.

14. Hill K, Vapattanawong P, Prasartkul P, Porapakkham Y, Lim SS, Lopez AD (2007). Epidemiologic transition interrupted: a reassessment of mortality trends in Thailand, 1980-2000. Int J Epidemiol 36(2): 374-384.

15. Hill K, Barker B, Vos T (2007). Excess Indigenous mortality: are Indigenous Australians more severely disadvantaged than other Indigenous populations? Int J Epidemiol 36(3): 580-589.

16. Merli MG (1998). Mortality in Vietnam, 1979-1989. Demography 35(3): 345-360.

17. Thomas KJA, Hill K (2007). Evaluating the performance of death distribution methods for estimating death registration completeness: applications to data from high-income countries. Paper presented in the session on Issues in the Measurement of Mortality. New York: Annual Meeting of the Population Association of America.

18. StataCorp LP (2008). Stata/SE 10.1 for Windows. College Station, TX: StataCorp.

19. United Nations Statistics Division (1989). Demographic yearbook 1989. New York: United Nations. http://unstats.un.org/unsd/demographic/products/dyb/dybsets/1989%20DYB.pdf. (accessed January 7, 2009).

20. Bhat PN (1990). Estimating transition probabilities of age misstatement. Demography 27(1): 149-163.

21. Preston SH, Lahiri S (1991). A short-cut method for estimating death registration completeness in destabilized populations. Math Popul Stud 3(1): 39-51.

22. Ezzati M, Friedman AB, Kulkarni SC, Murray CJ (2008). The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the United States. PLoS Med 5(4): e66.

23. Murray CJ, Kulkarni S, Ezzati M (2005). Eight Americas: new perspectives on U.S. health disparities. Am J Prev Med 29(5 Suppl 1): 4-10.

24. Murray CJ, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, Iandiorio TJ et al. (2006). Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. PLoS Med 3(9): e260.

25. Internal Revenue Service (2006). SOI tax stats - county-to-county migration data. Internal Revenue Service, United States Department of the Treasury. http://www.irs.gov/taxstats/indtaxstats/article/0,,id=96816,00.html (accessed January 7, 2009).

26. United Nations Statistics Division. Demographic yearbook. New York: United Nations. http://unstats.un.org/unsd/demographic/products/dyb/dyb2.htm (accessed January 7, 2009).

27. World Health Organization. Mortality database. Geneva: World Health Organization. http://www.who.int/whosis/mort/en/ (accessed January 7, 2009).

28. Minnesota Population Center. Integrated Public Use Microdata Series. IPUMS. USA. http://usa.ipums.org/usa/ (accessed January 7, 2009).

29. Lopez AD, Begg S, Bos E (2006). Demographic and epidemiological characteristics of major regions, 1990–2001. In: Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL, eds. Global burden of disease and risk factors. Oxford University Press and the World Bank: 17-44.

30. Thailand National Statistical Office (1997). Report on the 1995-1996 Survey of Population Change. Bangkok: Thailand National Statistical Office.

31. Chandra Sekar C, Deming WE (1949). On a method of estimating birth and death rates and the extent of registration. Journal of the American Statistical Association 44: 101-115.

32. Hook EB, Regal RR (1995). Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev 17(2): 243-264.

33. Yang G, Hu J, Rao KQ, Ma J, Rao C, Lopez AD (2005). Mortality registration and surveillance in China: History, current situation and challenges. Population Health Metrics 3(1): 3.

34. Eisele TP, Lindblade KA, Rosen DH, Odhiambo F, Vulule JM, Slutsker L (2003). Evaluating the completeness of demographic surveillance of children less than five years old in western Kenya: a capture-recapture approach. Am J Trop Med Hyg 69(1): 92-97.

35. Prasartkul P, Wapattanavong P (2009). Personal communication.

36. Preston SH, Elo IT (1999). Effects of age misreporting on mortality estimates at older ages. Population Studies 53(2): 165-177.

37. Vincent P (1951). La mortalité des vieillards. Population 6: 182-204.

Tables

Scenario	Crude Birth Rate per 1000	Life Expectancy at Birth	45q15	Population at Census 1 (in millions)	Population at Census 2 (in millions)	Intercensal Deaths (in millions)
1	7.55	75.17	0.106	15.0	15.3	1.48
2	8.83	72.47	0.144	12.0	13.1	1.13
3	8.77	73.01	0.135	12.6	13.7	1.19
4	7.78	74.54	0.115	14.6	15	1.42
5	8.42	72.88	0.138	12.4	13.5	1.18
6	9.09	71.85	0.154	11.2	12.3	1.05
7	8.78	72.80	0.139	11.6	12.4	1.10
8	8.58	72.91	0.139	11.7	12.5	1.11
9	7.62	75.00	0.109	14.3	14.9	1.41
10	7.74	74.88	0.110	14.2	14.9	1.41
11	8.97	72.35	0.148	11.1	12	1.06

Table 1: Mortality and Fertility Levels and Trends in the Simulations.

	Mean	Min	Max
Completeness of Census 1	0.95	0.90	1.00
Completeness of Census 2	0.95	0.90	1.00
Completeness of VR	0.65	0.30	1.00
β in Census 1	0.00	-0.07	0.07
σ_1^2 in Census 1	1.99	0.00	2.89
β in Census 2	0.00	-0.07	0.08
σ_1^2 in Census 2	1.99	0.00	2.86
β in VR	0.00	-0.06	0.07
σ_1^2 in VR	1.99	0.00	2.76

Notes:
VR = Vital Registration

Stochastic age-misreporting is captured as a random draw for each individual for each measurement from a normal distribution with mean zero and variance σ_1^2

Systematic age-misreporting is captured by the function $a_m=a_t+a_t*\beta$ where a_m is the misreported age, at is the true age, and β is drawn from a normal distribution.

Table 2: Simulation Measurement Model

	Years	Mean	Min	Max
Population Size	2000	401,245	100,224	9,519,338
Exponential Growth Rate	1990-2000	0.015	-0.013	0.103
Emmigration (total outmigrants)	1995-2000	22,711	547	1,190,823
Immigration (total inmigrants)	1995-2000	22,711	226	770,306
Net Migration	1995-2000	0	-420,571	193,489
Life Expectancy at Birth	1990-2000	78	70	82
45q15	1990-2000	0	0	0
Total Fertility Rate	2000	2.01	0	3.62

Table 3: Summary of Demographic Characteristics of U.S. Counties with Population Greater than 100,000 in 2000

	Simulations		US Counties		High-Income Countries		Average	
Age Trim	ARE	Rank	ARE	Rank	ARE	Rank	ARE	Rank
GGB								
30to80	0.091	18	0.126	3	0.048	1	0.088	7.3
25to80	0.091	20	0.125	2	0.048	2	0.088	8.0
35to80	0.091	19	0.128	6	0.049	4	0.089	9.7
20to80	0.092	22	0.126	4	0.048	3	0.089	9.7
15to80	0.092	27	0.127	5	0.049	5	0.089	12.3
…								…
25to45	0.647	78	0.358	68	0.167	64	0.391	70.0
10to35	0.210	59	0.454	75	0.279	77	0.314	70.3
15to40	0.393	72	0.387	69	0.232	73	0.338	71.3
15to35	0.310	68	0.451	74	0.318	78	0.360	73.3
20to40	0.520	76	0.407	72	0.233	74	0.387	74.0
SEG								
60to80	0.097	1	0.119	1	0.050	1	0.089	1.0
55to80	0.105	2	0.133	2	0.052	2	0.097	2.0
55to75	0.111	3	0.143	3	0.054	3	0.103	3.0
50to80	0.114	4	0.152	4	0.054	4	0.107	4.0
50to75	0.122	5	0.164	5	0.056	6	0.114	5.3
…								…
5to40	0.540	74	1.078	74	0.139	74	0.586	74.0
10to30	0.571	75	1.138	75	0.151	76	0.620	75.3
5to35	0.579	76	1.176	76	0.150	75	0.635	75.7
5to30	0.621	77	1.290	77	0.161	77	0.691	77.0
5to25	0.665	78	1.417	78	0.174	78	0.752	78.0
GGBSEG								
50to70	0.113	6	0.114	3	0.061	6	0.096	5.0
50to75	0.111	5	0.115	10	0.059	5	0.095	6.7
50to80	0.110	4	0.117	12	0.058	4	0.095	6.7
45to75	0.117	8	0.114	4	0.062	8	0.097	6.7
45to70	0.120	9	0.113	2	0.064	9	0.099	6.7
…								…
5to40	0.325	74	0.339	74	0.227	74	0.297	74.0
10to30	0.341	75	0.374	76	0.244	75	0.320	75.3
5to35	0.347	76	0.372	75	0.245	76	0.321	75.7
5to30	0.370	77	0.412	77	0.264	77	0.349	77.0
5to25	0.394	78	0.454	78	0.284	78	0.377	78.0

Table 4: Average Relative Error and Rank for the Top 5 and Worst 5 of 78 Possible Age-Trims of GGB, SEG, and GGBSEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

Variables	GGB est.	95% CI		t	SEG est.	95% CI		t	GGBSEG est.	95% CI		t
(a) Without Fixed Effects												
β in Census 1	-13.00	-13.03	-12.98	-1185.6	-9.81	-9.86	-9.76	-399.9	-15.22	-15.25	-15.19	-986.3
σ_1^2 in Census 1	-0.02	-0.03	-0.02	-27.4	0.00	0.00	0.01	1.3	-0.04	-0.04	-0.04	-31.8
β in Census 2	5.21	5.19	5.23	464.0	5.78	5.74	5.83	230.3	1.52	1.49	1.55	96.3
σ_1^2 in Census 2	0.02	0.02	0.02	24.6	0.02	0.02	0.02	10.0	0.02	0.02	0.03	19.3
β in VR	8.68	8.65	8.70	769.6	4.84	4.79	4.89	191.8	14.10	14.07	14.13	889.0
σ_1^2 in VR	0.00	0.00	0.01	4.7	-0.02	-0.03	-0.02	-12.4	0.02	0.01	0.02	12.2
Migrants per 1000	0.00	0.00	0.00	-237.8	0.00	0.00	0.00	284.1	0.00	0.00	0.00	-40.9
Constant	0.03	0.03	0.03	50.2	0.02	0.02	0.02	14.8	0.05	0.04	0.05	58.0
R^2	0.91				0.63				0.89			
RMSE	0.03				0.07				0.05			
(b) Including Fixed Effects for Population Scenario												
β in Census 1	-13.00	-13.02	-12.99	-1484.0	-9.81	-9.85	-9.78	-521.4	-15.22	-15.25	-15.19	-1069.5
σ_1^2 in Census 1	-0.02	-0.03	-0.02	-34.3	0.00	0.00	0.01	1.7	-0.04	-0.04	-0.04	-34.4
β in Census 2	5.21	5.19	5.23	580.8	5.78	5.75	5.82	300.3	1.52	1.49	1.55	104.4
σ_1^2 in Census 2	0.02	0.02	0.02	30.7	0.02	0.02	0.02	13.1	0.02	0.02	0.03	20.9
β in VR	8.68	8.66	8.69	963.3	4.84	4.80	4.87	250.1	14.10	14.07	14.13	964.0
σ_1^2 in VR	0.00	0.00	0.01	5.8	-0.02	-0.03	-0.02	-16.1	0.02	0.01	0.02	13.2
Migrants per 1000	0.00	0.00	0.00	-65.7	0.01	0.01	0.01	79.1	0.00	0.00	0.00	-12.8
Constant	0.03	0.03	0.03	36.4	0.02	0.02	0.02	12.3	0.03	0.03	0.04	28.6
R^2	0.94				0.78				0.90			
RMSE	0.03				0.06				0.04			

Notes: Stochastic age-misreporting is captured as a random draw for each individual from a normal distribution with mean zero and variance σ_1^2. Systematic age-misreporting is captured by the function $a_m=a_t+a_t*\beta$ where a_m is the misreported age, a_t is the true age, and β is drawn from a normal distribution.

Table 5: Coefficients from Regression of Error on Age Misreporting and Migration in the Simulations. This table shows the relationship between levels of age-misreporting and migration and error in relative completeness (RC) in the simulation environment both in the absence (a) and presence (b) of fixed effects indicating the combination of mortality, fertility and migration rates that define a population scenario. Error is calculated by subtracting true RC from estimated RC using the optimal variant in the simulated environment for each of the 3 families.

GGB

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
30to80	0.091	18	0.126	3	0.048	1	0.088	7.3
25to80	0.091	20	0.125	2	0.048	2	0.088	8.0
35to80	0.091	19	0.128	6	0.049	4	0.089	9.7
20to80	0.092	22	0.126	4	0.048	3	0.089	9.7
15to80	0.092	27	0.127	5	0.049	5	0.089	12.3
35to75	0.082	8	0.143	17	0.051	12	0.092	12.3
30to75	0.084	11	0.139	14	0.051	13	0.091	12.7
5to80	0.092	23	0.130	9	0.049	8	0.090	13.3
25to75	0.086	12	0.139	13	0.051	15	0.092	13.3
40to80	0.092	24	0.131	10	0.049	7	0.091	13.7
10to80	0.092	28	0.129	8	0.049	6	0.090	14.0
40to75	0.081	5	0.150	24	0.051	14	0.094	14.3
35to70	0.082	6	0.148	21	0.053	19	0.094	15.3
20to75	0.088	15	0.142	15	0.052	18	0.094	16.0
55to80	0.099	33	0.128	7	0.051	11	0.093	17.0
50to80	0.096	31	0.132	11	0.049	9	0.092	17.0
45to80	0.093	29	0.133	12	0.049	10	0.092	17.0
15to75	0.090	17	0.143	16	0.053	20	0.095	17.7
45to75	0.082	7	0.157	29	0.052	17	0.097	17.7
40to70	0.079	4	0.158	31	0.053	21	0.097	18.7
30to70	0.086	14	0.146	19	0.053	23	0.095	18.7
50to75	0.083	10	0.160	34	0.052	16	0.098	20.0
60to80	0.104	38	0.122	1	0.054	24	0.093	21.0
45to70	0.078	1	0.170	40	0.053	22	0.100	21.0
5to75	0.091	21	0.147	20	0.055	27	0.098	22.7
10to75	0.092	26	0.144	18	0.054	25	0.097	23.0
55to75	0.086	13	0.159	33	0.054	26	0.100	24.0
35to65	0.090	16	0.149	22	0.064	35	0.101	24.3
40to65	0.082	9	0.154	28	0.065	36	0.101	24.3
50to70	0.078	2	0.179	45	0.056	29	0.105	25.3
25to70	0.092	25	0.151	25	0.055	28	0.099	26.0
45to65	0.078	3	0.170	39	0.067	38	0.105	26.7
20to70	0.097	32	0.158	30	0.057	30	0.104	30.7
30to65	0.099	34	0.154	27	0.064	34	0.106	31.7
15to70	0.101	35	0.159	32	0.059	31	0.107	32.7
40to60	0.095	30	0.150	23	0.085	46	0.110	33.0
10to70	0.103	37	0.161	35	0.061	32	0.108	34.7

Appendix Table 1a: Average Relative Error and Rank for All Possible Age-Trims of GGB in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

GGB

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
5to70	0.102	36	0.164	36	0.063	33	0.110	35.0
35to60	0.111	40	0.152	26	0.075	41	0.113	35.7
25to65	0.110	39	0.165	37	0.067	37	0.114	37.7
20to65	0.120	41	0.176	43	0.070	39	0.122	41.0
15to65	0.126	43	0.176	42	0.072	40	0.125	41.7
30to60	0.130	45	0.168	38	0.077	43	0.125	42.0
10to65	0.129	44	0.178	44	0.075	42	0.127	43.3
5to65	0.124	42	0.182	46	0.077	44	0.128	44.0
25to60	0.149	46	0.187	47	0.080	45	0.139	46.0
35to55	0.156	48	0.176	41	0.094	51	0.142	46.7
20to60	0.164	50	0.202	49	0.086	47	0.151	48.7
15to60	0.173	53	0.200	48	0.090	48	0.154	49.7
10to60	0.172	52	0.203	50	0.093	50	0.156	50.7
5to60	0.158	49	0.208	52	0.096	52	0.154	51.0
30to55	0.192	56	0.206	51	0.091	49	0.163	52.0
25to55	0.225	61	0.230	53	0.098	53	0.184	55.7
5to55	0.194	57	0.250	57	0.117	58	0.187	57.3
20to55	0.248	63	0.245	56	0.106	54	0.200	57.7
10to55	0.236	62	0.243	55	0.114	57	0.198	58.0
15to55	0.252	64	0.239	54	0.112	56	0.201	58.0
30to50	0.337	70	0.270	58	0.112	55	0.240	61.0
5to50	0.215	60	0.297	62	0.143	63	0.218	61.7
10to50	0.307	67	0.289	60	0.142	62	0.246	63.0
5to45	0.203	58	0.346	66	0.176	66	0.242	63.3
15to50	0.377	71	0.285	59	0.140	61	0.267	63.7
25to50	0.397	73	0.290	61	0.118	59	0.268	64.3
5to40	0.176	54	0.392	70	0.214	70	0.261	64.7
20to50	0.414	74	0.298	63	0.132	60	0.282	65.7
5to35	0.167	51	0.438	73	0.248	75	0.284	66.3
10to30	0.150	47	0.545	78	0.270	76	0.322	67.0
10to45	0.327	69	0.340	65	0.179	67	0.282	67.0
5to30	0.189	55	0.479	77	0.228	72	0.299	68.0
15to45	0.463	75	0.333	64	0.180	68	0.325	69.0

Appendix Table 1a: Average Relative Error and Rank for All Possible Age-Trims of GGB in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

GGB

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
10to40	0.274	66	0.393	71	0.224	71	0.297	69.3
20to45	0.599	77	0.351	67	0.174	65	0.375	69.7
5to25	0.252	65	0.459	76	0.207	69	0.306	70.0
25to45	0.647	78	0.358	68	0.167	64	0.391	70.0
10to35	0.210	59	0.454	75	0.279	77	0.314	70.3
15to40	0.393	72	0.387	69	0.232	73	0.338	71.3
15to35	0.310	68	0.451	74	0.318	78	0.360	73.3
20to40	0.520	76	0.407	72	0.233	74	0.387	74.0

Appendix Table 1a: Average Relative Error and Rank for All Possible Age-Trims of GGB in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

SEG

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
60to80	0.097	1	0.119	1	0.050	1	0.089	1.0
55to80	0.105	2	0.133	2	0.052	2	0.097	2.0
55to75	0.111	3	0.143	3	0.054	3	0.103	3.0
50to80	0.114	4	0.152	4	0.054	4	0.107	4.0
50to75	0.122	5	0.164	5	0.056	6	0.114	5.3
45to80	0.126	6	0.174	6	0.056	5	0.119	5.7
50to70	0.131	7	0.180	7	0.059	9	0.123	7.7
45to75	0.135	8	0.188	8	0.058	8	0.127	8.0
40to80	0.141	9	0.199	9	0.058	7	0.132	8.3
45to70	0.145	10	0.206	10	0.060	12	0.137	10.7
40to75	0.151	11	0.215	11	0.060	10	0.142	10.7
35to80	0.158	13	0.229	12	0.060	11	0.149	12.0
45to65	0.156	12	0.229	13	0.063	15	0.149	13.3
40to70	0.162	14	0.235	14	0.062	13	0.153	13.7
35to75	0.169	15	0.247	15	0.063	14	0.160	14.7
40to65	0.174	16	0.259	16	0.065	17	0.166	16.3
30to80	0.180	17	0.268	17	0.064	16	0.171	16.7
35to70	0.182	18	0.269	18	0.065	18	0.172	18.0
40to60	0.188	19	0.287	19	0.067	20	0.181	19.3
30to75	0.192	20	0.289	20	0.066	19	0.182	19.7
35to65	0.195	21	0.295	21	0.068	21	0.186	21.0
25to80	0.205	22	0.324	23	0.068	22	0.199	22.3
30to70	0.206	23	0.313	22	0.069	23	0.196	22.7
35to60	0.211	24	0.324	24	0.071	24	0.202	24.0
25to75	0.219	25	0.348	26	0.071	25	0.213	25.3
30to65	0.221	26	0.342	25	0.072	26	0.212	25.7
35to55	0.228	27	0.356	27	0.073	28	0.219	27.3
20to80	0.234	28	0.388	30	0.073	27	0.232	28.3
25to70	0.234	29	0.376	29	0.074	29	0.228	29.0
30to60	0.238	30	0.374	28	0.075	30	0.229	29.3
25to65	0.251	32	0.408	31	0.077	32	0.245	31.7
20to75	0.249	31	0.416	33	0.076	31	0.247	31.7
30to55	0.257	33	0.409	32	0.078	33	0.248	32.7
15to80	0.263	34	0.450	37	0.079	34	0.264	35.0
20to70	0.265	35	0.448	36	0.080	35	0.264	35.3

Appendix Table 1b: Average Relative Error and Rank for All Possible Age-Trims of SEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

SEG

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
25to60	0.270	36	0.445	34	0.081	36	0.265	35.3
30to50	0.279	38	0.447	35	0.081	37	0.269	36.7
15to75	0.279	37	0.481	38	0.083	38	0.281	37.7
20to65	0.284	39	0.484	39	0.084	39	0.284	39.0
25to55	0.291	40	0.485	40	0.084	40	0.287	40.0
10to80	0.293	41	0.522	42	0.086	41	0.300	41.3
15to70	0.297	42	0.516	41	0.087	42	0.300	41.7
20to60	0.304	43	0.525	43	0.087	43	0.306	43.0
25to50	0.315	45	0.530	44	0.088	44	0.311	44.3
10to75	0.310	44	0.556	46	0.090	45	0.319	45.0
15to65	0.317	46	0.555	45	0.091	46	0.321	45.7
20to55	0.327	47	0.571	47	0.091	47	0.330	47.0
25to45	0.343	51	0.580	48	0.092	48	0.339	49.0
5to80	0.328	48	0.615	51	0.093	49	0.345	49.3
10to70	0.329	49	0.594	49	0.094	50	0.339	49.3
15to60	0.339	50	0.600	50	0.095	51	0.345	50.3
20to50	0.353	54	0.622	52	0.096	52	0.357	52.7
10to65	0.350	53	0.637	53	0.099	54	0.362	53.3
5to75	0.347	52	0.653	55	0.097	53	0.366	53.3
15to55	0.363	55	0.649	54	0.100	55	0.371	54.7
20to45	0.383	58	0.680	56	0.101	56	0.388	56.7
5to70	0.367	56	0.695	58	0.102	57	0.388	57.0
10to60	0.373	57	0.685	57	0.103	58	0.387	57.3
15to50	0.391	60	0.704	59	0.105	59	0.400	59.3
5to65	0.390	59	0.743	61	0.107	60	0.413	60.0
10to55	0.398	61	0.738	60	0.108	61	0.415	60.7
20to40	0.416	63	0.750	62	0.108	62	0.425	62.3
15to45	0.422	64	0.766	63	0.111	63	0.433	63.3
5to60	0.414	62	0.797	64	0.112	64	0.441	63.3
10to50	0.427	65	0.797	65	0.114	65	0.446	65.0
5to55	0.441	66	0.855	67	0.117	66	0.471	66.3
15to40	0.456	67	0.838	66	0.119	67	0.471	66.7
10to45	0.458	68	0.863	68	0.121	68	0.481	68.0

Appendix Table 1b: Average Relative Error and Rank for All Possible Age-Trims of SEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

SEG

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
5to50	0.471	69	0.921	69	0.123	69	0.505	69.0
10to40	0.493	70	0.940	71	0.129	70	0.521	70.3
15to35	0.494	71	0.925	70	0.129	71	0.516	70.7
5to45	0.504	72	0.994	72	0.131	72	0.543	72.0
10to35	0.531	73	1.031	73	0.139	73	0.567	73.0
5to40	0.540	74	1.078	74	0.139	74	0.586	74.0
10to30	0.571	75	1.138	75	0.151	76	0.620	75.3
5to35	0.579	76	1.176	76	0.150	75	0.635	75.7
5to30	0.621	77	1.290	77	0.161	77	0.691	77.0
5to25	0.665	78	1.417	78	0.174	78	0.752	78.0
10to40	0.274	66	0.393	71	0.224	71	0.297	69.3
20to45	0.599	77	0.351	67	0.174	65	0.375	69.7
5to25	0.252	65	0.459	76	0.207	69	0.306	70.0
25to45	0.647	78	0.358	68	0.167	64	0.391	70.0
10to35	0.210	59	0.454	75	0.279	77	0.314	70.3
15to40	0.393	72	0.387	69	0.232	73	0.338	71.3
15to35	0.310	68	0.451	74	0.318	78	0.360	73.3
20to40	0.520	76	0.407	72	0.233	74	0.387	74.0

Appendix Table 1b: Average Relative Error and Rank for All Possible Age-Trims of SEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

GGBSEG

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
50to70	0.113	6	0.114	3	0.061	6	0.096	5.0
50to75	0.111	5	0.115	10	0.059	5	0.095	6.7
50to80	0.110	4	0.117	12	0.058	4	0.095	6.7
45to75	0.117	8	0.114	4	0.062	8	0.097	6.7
45to70	0.120	9	0.113	2	0.064	9	0.099	6.7
55to75	0.105	2	0.118	16	0.055	3	0.093	7.0
55to80	0.105	3	0.119	17	0.055	2	0.093	7.3
60to80	0.102	1	0.123	20	0.053	1	0.092	7.3
45to80	0.115	7	0.115	9	0.061	7	0.097	7.7
45to65	0.125	12	0.112	1	0.069	12	0.102	8.3
40to75	0.124	11	0.114	6	0.067	11	0.102	9.3
40to80	0.122	10	0.115	8	0.065	10	0.101	9.3
40to70	0.128	13	0.114	5	0.070	13	0.104	10.3
40to65	0.134	16	0.115	7	0.075	16	0.108	13.0
35to80	0.129	14	0.117	13	0.071	14	0.106	13.7
35to75	0.132	15	0.117	14	0.074	15	0.108	14.7
40to60	0.141	19	0.116	11	0.081	19	0.113	16.3
35to70	0.137	17	0.118	15	0.078	17	0.111	16.3
30to80	0.138	18	0.122	19	0.080	18	0.113	18.3
35to65	0.143	21	0.120	18	0.083	21	0.115	20.0
30to75	0.142	20	0.124	22	0.083	20	0.116	20.7
30to70	0.148	22	0.126	23	0.087	22	0.120	22.3
35to60	0.151	24	0.123	21	0.090	24	0.121	23.0
25to80	0.150	23	0.134	26	0.089	23	0.124	24.0
30to65	0.155	26	0.129	25	0.094	26	0.126	25.7
25to75	0.154	25	0.137	28	0.093	25	0.128	26.0
35to55	0.160	27	0.128	24	0.098	27	0.129	26.0
25to70	0.161	28	0.141	30	0.098	28	0.133	28.7
30to60	0.164	30	0.134	27	0.101	30	0.133	29.0
20to80	0.163	29	0.149	32	0.100	29	0.137	30.0
25to65	0.169	32	0.146	31	0.105	32	0.140	31.7
30to55	0.174	33	0.141	29	0.110	33	0.142	31.7
20to75	0.168	31	0.154	35	0.104	31	0.142	32.3

Appendix Table 1c: Average Relative Error and Rank for All Possible Age-Trims of GGBSEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

GGBSEG

Age Trim	Simulations ARE	Simulations Rank	US Counties ARE	US Counties Rank	High-Income Countries ARE	High-Income Countries Rank	Average ARE	Average Rank
20to70	0.176	34	0.160	36	0.111	34	0.149	34.7
25to60	0.179	36	0.153	34	0.114	36	0.149	35.3
15to80	0.177	35	0.164	38	0.112	35	0.151	36.0
30to50	0.186	39	0.150	33	0.118	39	0.151	37.0
15to75	0.183	37	0.170	40	0.117	37	0.157	38.0
20to65	0.185	38	0.167	39	0.118	38	0.157	38.3
25to55	0.190	40	0.163	37	0.123	40	0.159	39.0
15to70	0.191	42	0.177	43	0.124	41	0.164	42.0
10to80	0.191	41	0.181	44	0.125	42	0.166	42.3
20to60	0.196	43	0.176	42	0.127	43	0.166	42.7
25to50	0.203	46	0.175	41	0.132	46	0.170	44.3
15to65	0.201	45	0.186	45	0.132	45	0.173	45.0
10to75	0.198	44	0.188	47	0.130	44	0.172	45.0
20to55	0.208	48	0.188	46	0.137	47	0.178	47.0
10to70	0.207	47	0.197	50	0.138	48	0.181	48.3
5to80	0.211	49	0.202	51	0.139	49	0.184	49.7
15to60	0.213	50	0.196	49	0.141	50	0.184	49.7
25to45	0.219	52	0.191	48	0.143	51	0.184	50.3
10to65	0.218	51	0.207	53	0.147	53	0.190	52.3
5to75	0.219	53	0.211	55	0.145	52	0.192	53.3
20to50	0.223	54	0.203	52	0.147	54	0.191	53.3
15to55	0.227	55	0.210	54	0.152	55	0.196	54.7
5to70	0.229	56	0.221	57	0.153	56	0.201	56.3
10to60	0.230	57	0.219	56	0.157	57	0.202	56.7
20to45	0.239	58	0.223	58	0.158	58	0.207	58.0
5to65	0.240	59	0.233	60	0.163	59	0.212	59.3
15to50	0.242	60	0.227	59	0.163	60	0.210	59.7
10to55	0.244	61	0.234	61	0.168	61	0.216	61.0
5to60	0.254	62	0.247	62	0.173	63	0.225	62.3
20to40	0.258	63	0.248	63	0.172	62	0.226	62.7
15to45	0.259	64	0.248	64	0.175	64	0.227	64.0
10to50	0.260	65	0.253	65	0.179	65	0.231	65.0
5to55	0.269	66	0.264	66	0.185	66	0.239	66.0
15to40	0.279	68	0.274	67	0.190	67	0.247	67.3
10to45	0.278	67	0.276	68	0.192	68	0.249	67.7

Appendix Table 1c: Average Relative Error and Rank for All Possible Age-Trims of GGBSEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

GGBSEG

	Simulations		US Counties		High-Income Countries		Average	
Age Trim	ARE	Rank	ARE	Rank	ARE	Rank	ARE	Rank
5to50	0.285	69	0.285	69	0.198	69	0.256	69.0
10to40	0.298	70	0.303	70	0.208	71	0.269	70.3
15to35	0.300	71	0.306	71	0.207	70	0.271	70.7
5to45	0.304	72	0.310	72	0.211	72	0.275	72.0
10to35	0.319	73	0.335	73	0.225	73	0.293	73.0
5to40	0.325	74	0.339	74	0.227	74	0.297	74.0
10to30	0.341	75	0.374	76	0.244	75	0.320	75.3
5to35	0.347	76	0.372	75	0.245	76	0.321	75.7
5to30	0.370	77	0.412	77	0.264	77	0.349	77.0
5to25	0.394	78	0.454	78	0.284	78	0.377	78.0
10to40	0.274	66	0.393	71	0.224	71	0.297	69.3
20to45	0.599	77	0.351	67	0.174	65	0.375	69.7
5to25	0.252	65	0.459	76	0.207	69	0.306	70.0
25to45	0.647	78	0.358	68	0.167	64	0.391	70.0
10to35	0.210	59	0.454	75	0.279	77	0.314	70.3
15to40	0.393	72	0.387	69	0.232	73	0.338	71.3
15to35	0.310	68	0.451	74	0.318	78	0.360	73.3
20to40	0.520	76	0.407	72	0.233	74	0.387	74.0

Appendix Table 1c: Average Relative Error and Rank for All Possible Age-Trims of GGBSEG in the Simulations, U.S. Counties and High-Income Countries Sorted by Average Rank

Figure Legends

Figure 1: Three Families of Death Distribution Methods

Figure 2a: Simulated population model. This schematic describes the evolution of the simulated population, where P_0 is the probability of remaining in the sample, P_m is the probability of dying in the year given an age-specific probability ψ of dying in a single day (ζ is the fraction of time spent in the year in age group x1), $P_\varepsilon(a,t)$ is the probability of migrating at age a and time t, and $P_b(a,t)$ is the probability of giving birth at age a and time t and only applies to the reproductive age groups.

Figure 2b: The effect of fertility and mortality evolution over time on the population age-structure with no migration. q is the probability of dying per 1000.

Figure 2c: The age pattern of in- and out- migration used to model migration in the simulations. This age pattern is based on the average of a geographically diverse selection of countries with complete migration data as reported in the 1989 Demographic Yearbook.

Figure 3: Estimated relative completeness versus true relative completeness in the simulations, U.S. counties with a population greater than 100,000 and large high-income countries for the three methods. The true completeness for the counties is artificially offset from 1 in order to better distinguish it graphically from high-income countries.

Figure 4: Application of optimal Death Distribution Methods to 8 countries.

Figure 1

Generalized Growth Balance (GGB)

Hill, 1987

Data required:

- age and sex distribution of population at two time points, census 1 and census 2
- average annual deaths during the intercensal period

Base concept: the mathematical relationship of the demographic balancing equation:

Birth rate = growth rate + death rate

The slope and intercept of the modeled equation (plotting birth rate versus death rate) along with observed growth rates can be used to obtain the relative coverage of census 1 to census 2 as well as the relative completeness of death registration to census coverage.

Assumptions:

- No migration
- No age-misreporting
- Completeness constant by age

Synthetic Extinct Generations (SEG)

Bennett & Horiuchi, 1981, 1984

Data required:

- age and sex distribution of population at two time points, census 1 and census 2
- average annual deaths during the intercensal period

Base concept: the number of people age x at time 0 is equal to the number of deaths age x in year 0, plus deaths age x+1 in year 1, plus deaths age x+2 in year 2 and so on until the entire cohort is extinct. Use intercensal age-specific growth rates and current deaths at older ages to estimate future cohort deaths.

By comparing estimated future cohort deaths to current cohort size, the completeness of death registration can be obtained.

Assumptions:

- No migration
- No age-misreporting
- Completeness constant by age
- Coverage constant across two censuses

Hybrid (GGB-SEG)

Hill & Choi, 2004

Data required:

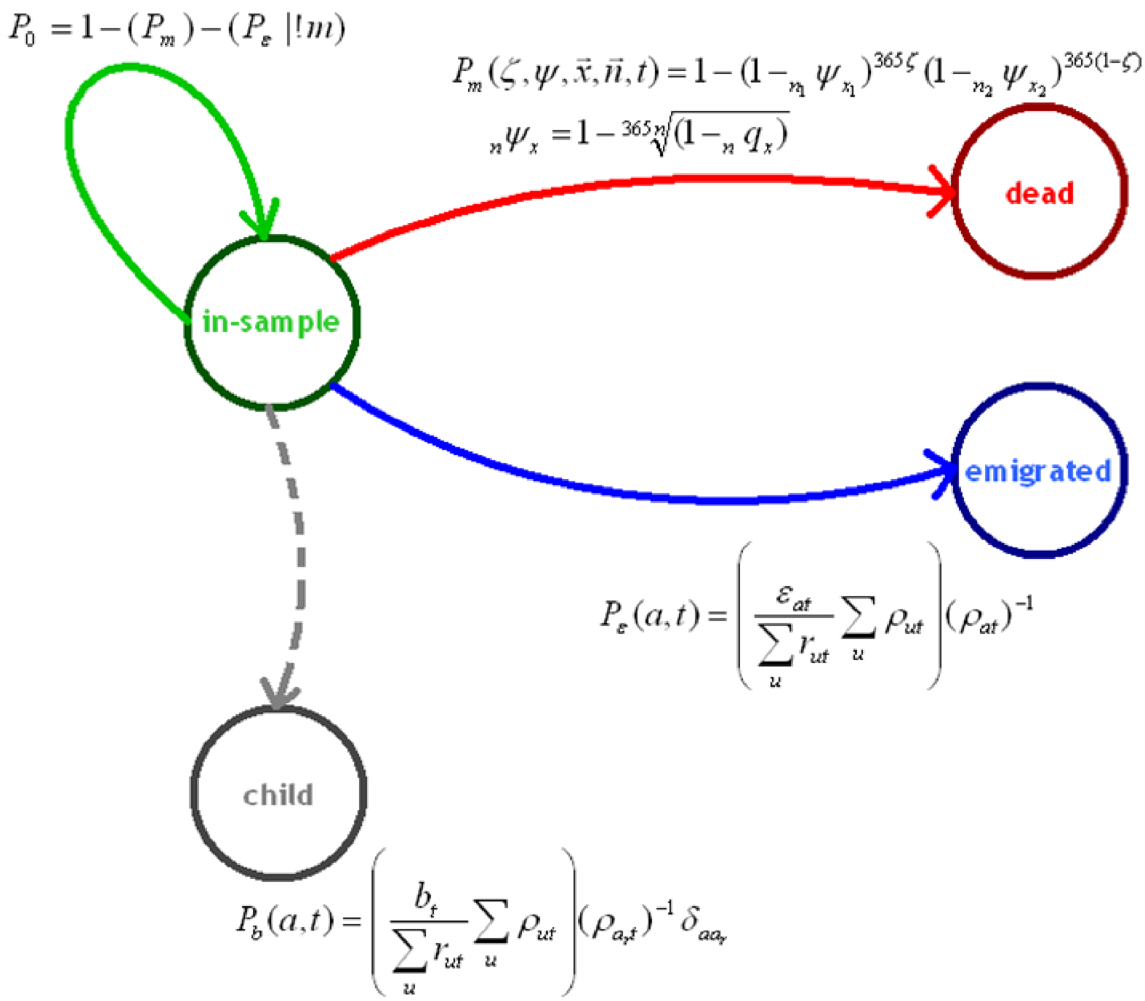
- age and sex distribution of population at two time points, census 1 and census 2
- average annual deaths during the intercensal period

Base concept: Use GGB to estimate coverage of census 2 relative to census 1 and use this to adjust populations prior to use in SEG method. This allows relaxation of the SEG assumption of constant coverage between censuses.

Assumptions:

- No migration
- No age-misreporting
- Completeness constant by age
- Relative coverage of censuses constant by age

Figure 2a



$P_0 = 1-(P_m)-(P_\varepsilon \,|!m)$
$P_m(\zeta,\psi,\vec{x},\vec{n},t) = 1-(1-{}_{n_1}\psi_{x_1})^{365\zeta}(1-{}_{n_2}\psi_{x_2})^{365(1-\zeta)}$
${}_n\psi_x = 1-\sqrt[365n]{(1-{}_n q_x)}$
dead
in-sample
emigrated
$P_\varepsilon(a,t) = \left(\frac{\varepsilon_{at}}{\sum_u r_{ut}}\sum_u \rho_{ut}\right)(\rho_{at})^{-1}$
child
$P_b(a,t) = \left(\frac{b_t}{\sum_u r_{ut}}\sum_u \rho_{ut}\right)(\rho_{a_f t})^{-1}\delta_{aa_f}$

Figure 2b

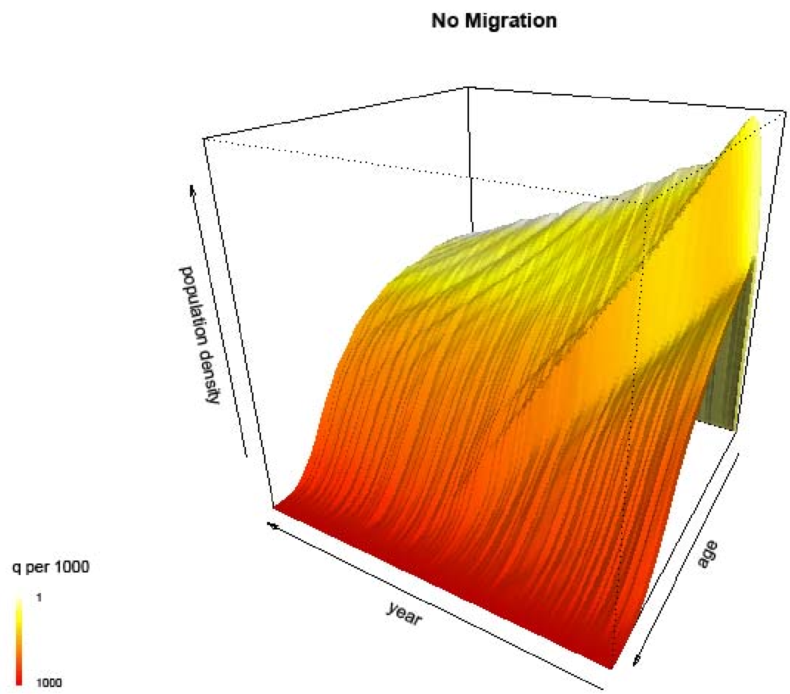


Figure 2c

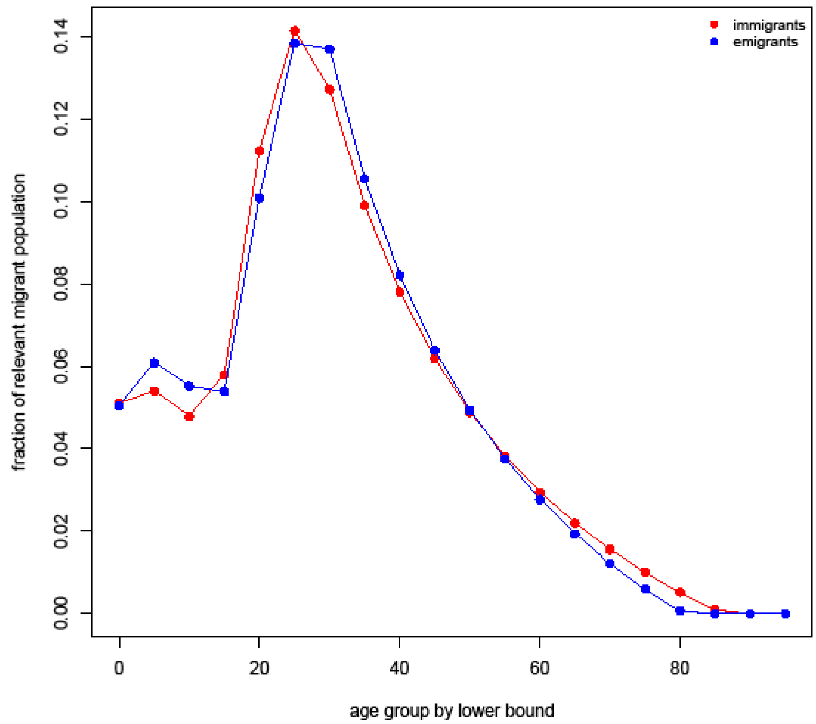

immigrants
emigrants
fraction of relevant migrant population
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0
20
40
60
80
age group by lower bound

Figure 3

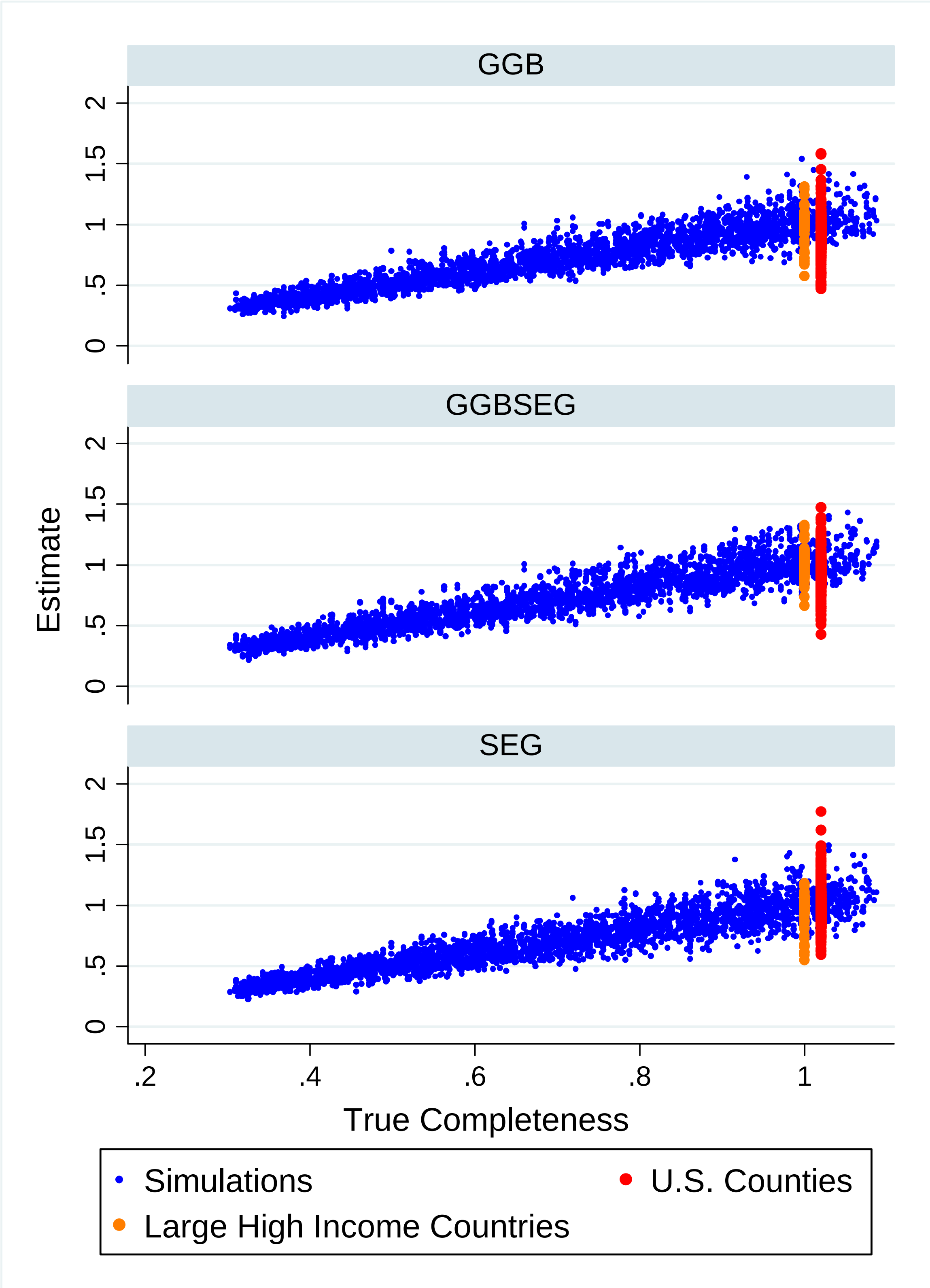
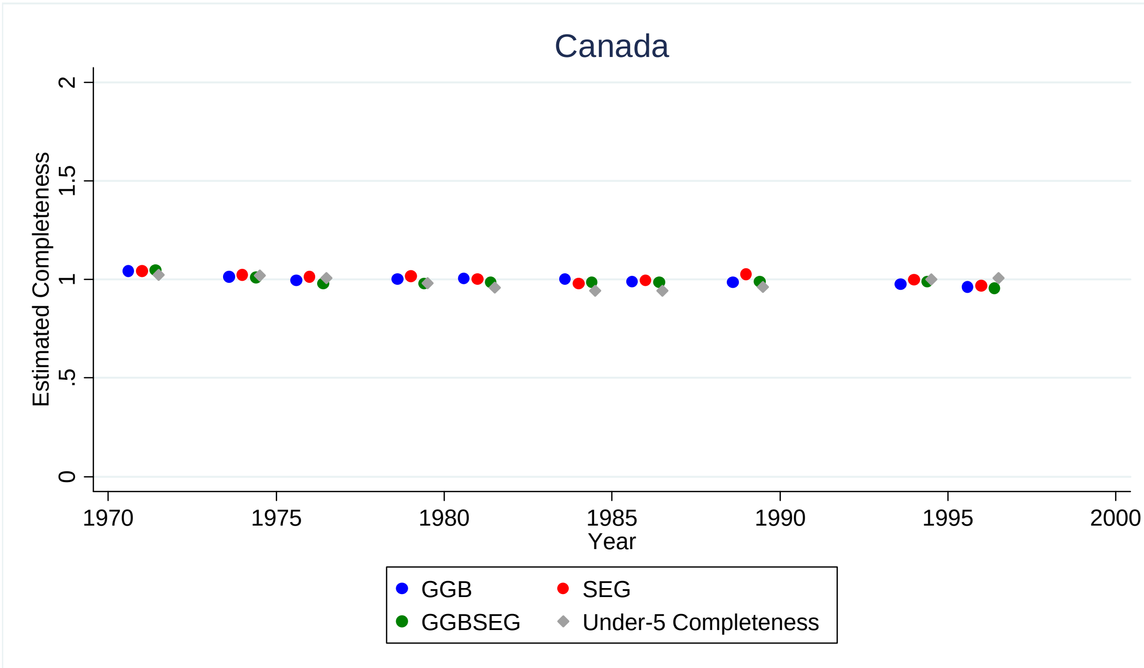
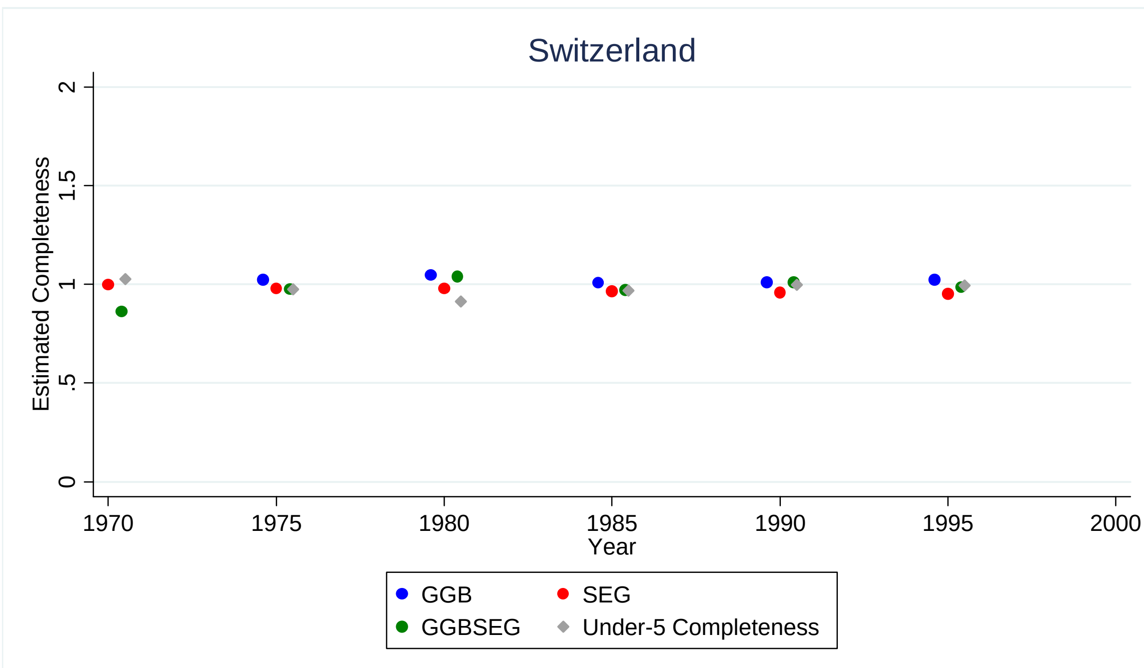
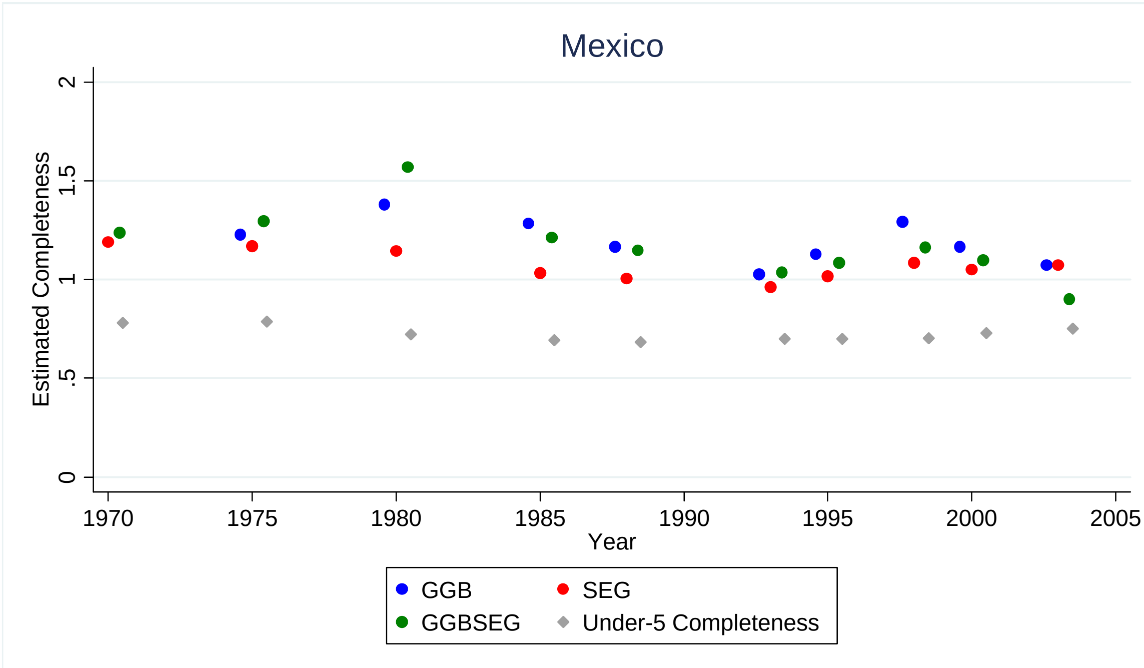
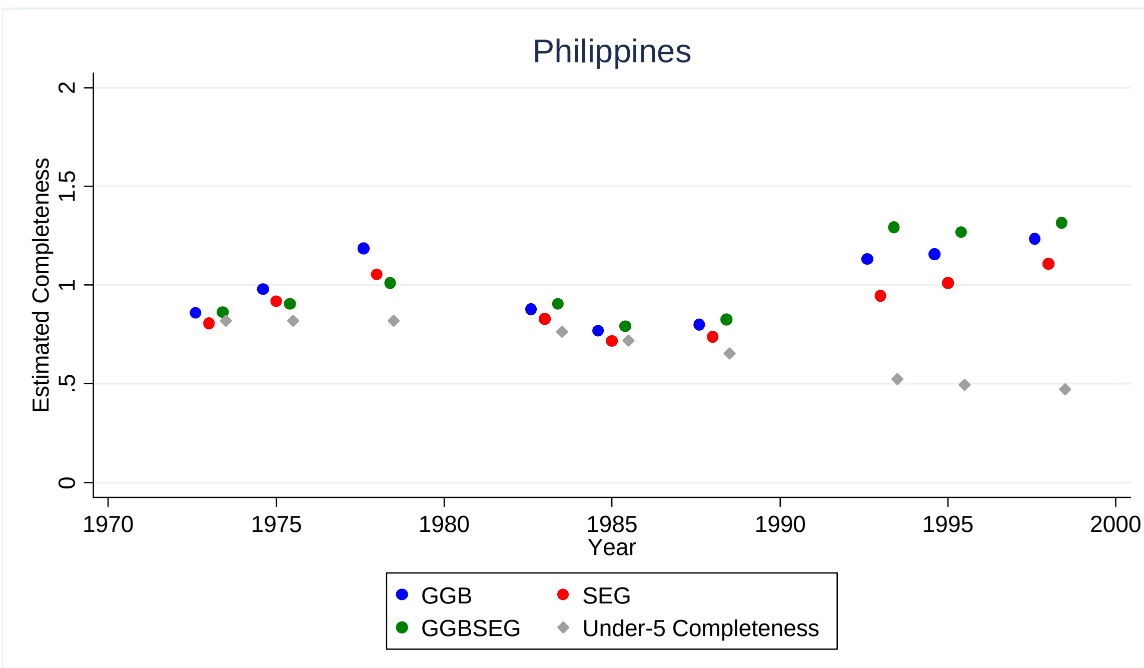

GGB
GGBSEG
SEG
Estimate
0
.5
1
1.5
2
.2
.4
.6
.8
1
True Completeness
Simulations
U.S. Counties
Large High Income Countries

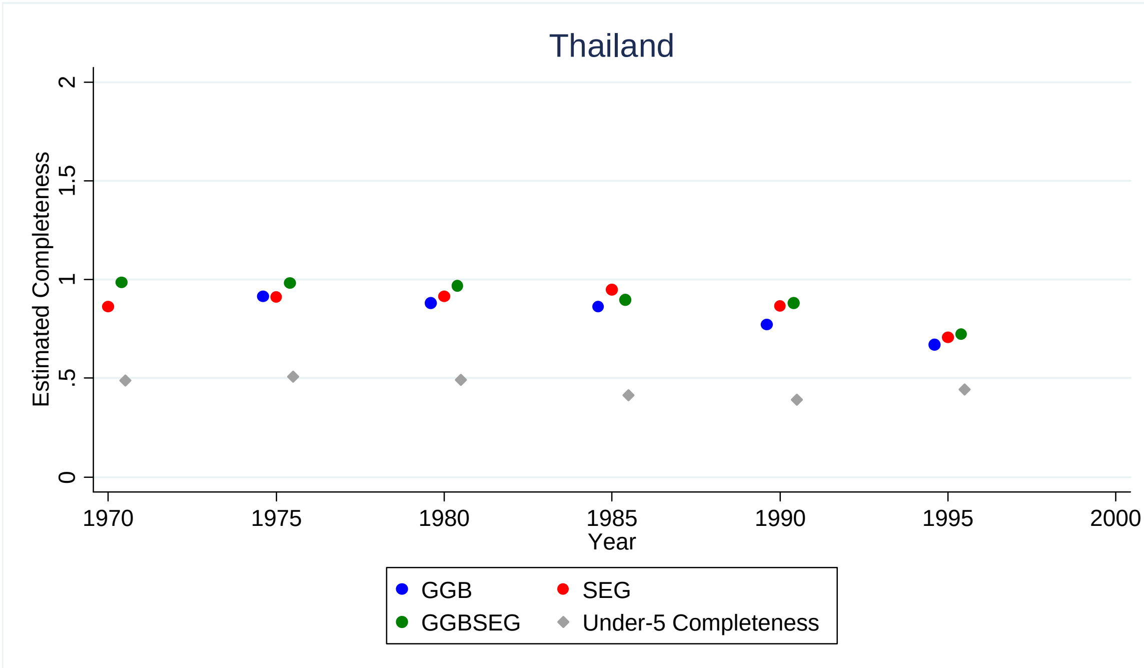
Figure 4

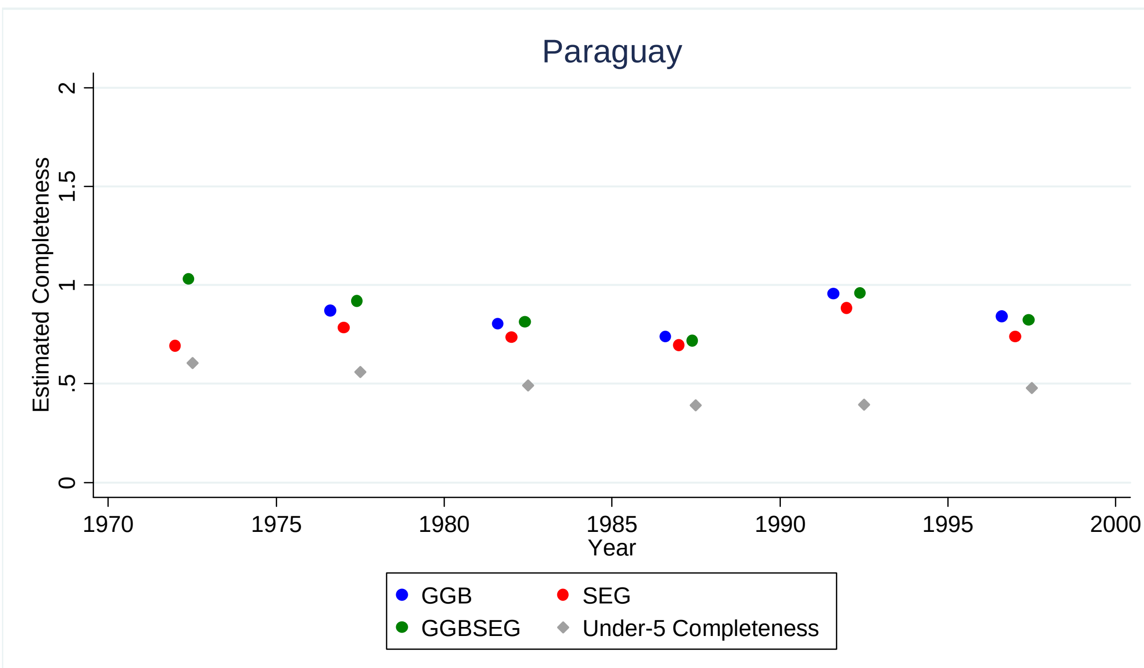

Canada
Estimated Completeness
0
.5
1
1.5
2
1970
1975
1980
1985
1990
1995
2000
Year
GGB
SEG
GGBSEG
Under-5 Completeness

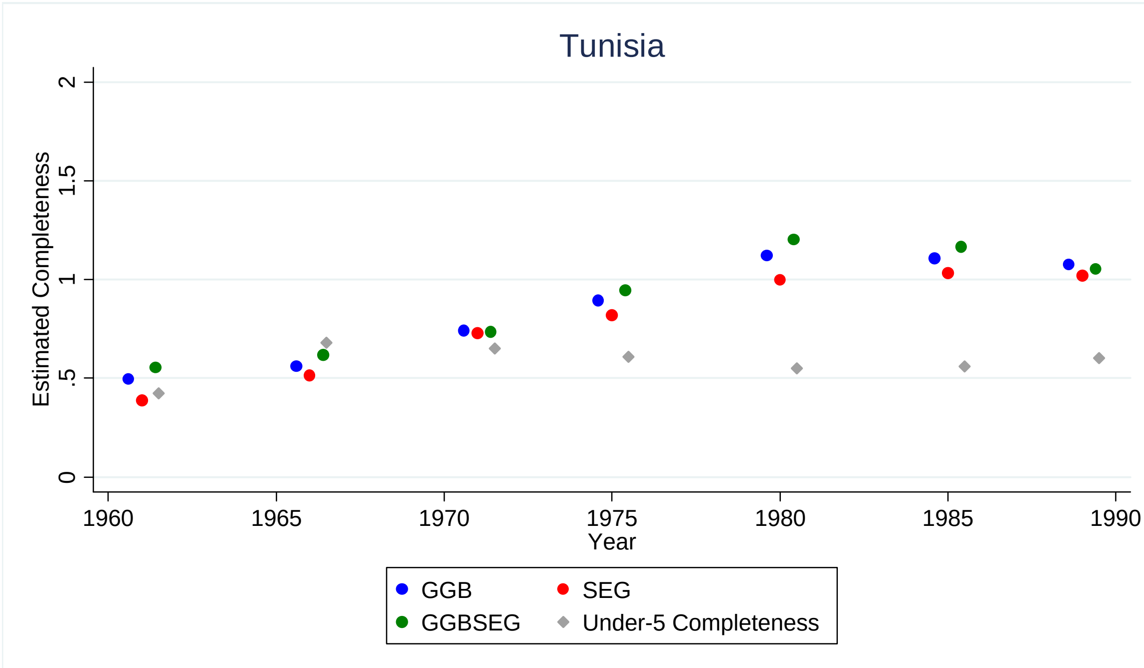

Switzerland
Estimated Completeness
0
.5
1
1.5
2
1970
1975
1980
1985
1990
1995
2000
Year
GGB
SEG
GGBSEG
Under-5 Completeness

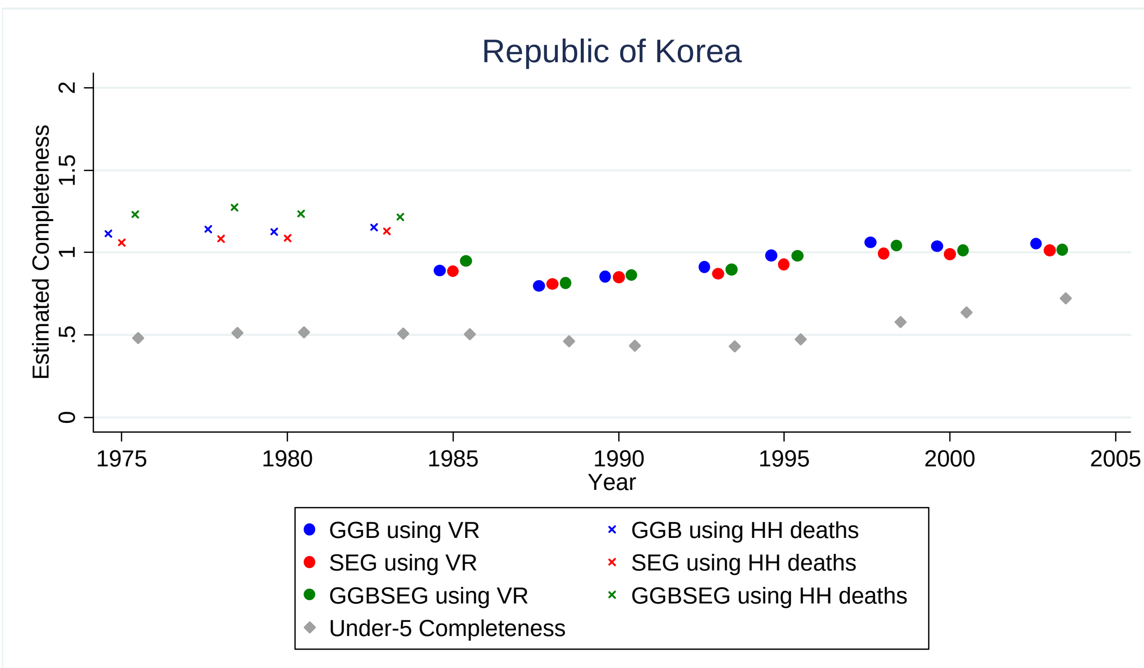

Mexico
Estimated Completeness
0
.5
1
1.5
2
1970
1975
1980
1985
1990
1995
2000
2005
Year
GGB
SEG
GGBSEG
Under-5 Completeness


Philippines
Estimated Completeness
0
.5
1
1.5
2
1970
1975
1980
1985
1990
1995
2000
Year
GGB
SEG
GGBSEG
Under-5 Completeness


Thailand
Estimated Completeness
0
.5
1
1.5
2
1970
1975
1980
1985
1990
1995
2000
Year
GGB
SEG
GGBSEG
Under-5 Completeness


Paraguay
Estimated Completeness
0
.5
1
1.5
2
1970
1975
1980
1985
1990
1995
2000
Year
GGB
SEG
GGBSEG
Under-5 Completeness


Tunisia
Estimated Completeness
0
.5
1
1.5
2
1960
1965
1970
1975
1980
1985
1990
Year
GGB
SEG
GGBSEG
Under-5 Completeness


Republic of Korea
Estimated Completeness
0
.5
1
1.5
2
1975
1980
1985
1990
1995
2000
2005
Year
GGB using VR
GGB using HH deaths
SEG using VR
SEG using HH deaths
GGBSEG using VR
GGBSEG using HH deaths
Under-5 Completeness