BACKGROUND NOTE ON METHODOLOGY FOR UNDER-FIVE MORTALITY ESTIMATION

Background and Inter-Agency work

Global estimates on under-five mortality are produced on a yearly basis by the **Inter-agency Group for Child Mortality Estimation**, which includes UNICEF, WHO, The World Bank, UN Population Division, Harvard University and others. The Interagency group was created in 2004 to share data and ensure consistency among the estimates previously produced separately by UNICEF, WHO and The World Bank. Since then the group has branched out to address common concerns in improving child mortality estimation. Membership has also expanded to include the UN Population Division and experts from universities and research institutes, followed by the US Bureau of the Census, ECLAC and the Demographic and Health Surveys group.

Since its inception, the Group has been driven by a common desire to improve its estimates, refine its working methods and expand its membership to include more experts.

A major database of mortality source data is being developed and builds on the databases previously maintained by UNICEF and WHO.

The Inter-agency Group updates the under-five mortality estimates on a yearly basis, undertaking a detailed review of all new available data points. At times, the review results in adjustments to previously reported estimates. Therefore, estimates published in consecutive editions of the agency's reports (e.g. UNICEF's *The State of the World's Children*) may not be comparable and should not be used for analyzing mortality trends over time.

Methodology

At country level, under-five mortality can be measured using a number of different methods, including registration of births and deaths via vital registration systems, national population censuses and/or data collected via household surveys. When vital registration systems are of good quality, the under-five mortality can be easily estimated. However, in the developing world most countries do not have well-functioning vital registration systems which can generate nationally representative estimates. Under-five mortality estimates can also be derived from household survey data using direct or indirect methods, including the MICS and DHS surveys.

These national level estimates are compiled by UNICEF and WHO as part of the work of the Inter-agency Child Mortality Estimation Group and are then plotted. The methodology used to derive the final estimate depends on the type of available data. For the countries with data from household surveys and censuses, most of the best estimates were derived by fitting a regression line to the relationship between infant or under 5 mortality rates and their reference dates using weighted least squares. A curve is then fitted to these points and extrapolated to a common reference year. The method aims to provide a transparent and largely objective way of fitting a smoothed trend to a set of observations, and of extrapolating the trend to cover the period from 1960 to the present.

The full technical details of this methodology, which is used by the Inter-agency Group for Child Mortality Estimation, can be found in ***Trends in Child Mortality in the Developing World: 1960-1990***[1] **by K. Hill et al (see annex 1). For an example of the process in practice, please see annex 2 on Egypt.**

Global estimates of the number of under-five deaths

The global estimates of the total number of deaths under the age of five are obtained by multiplying the estimated under-five mortality rate by the total number of births estimated by the UN Population Division for each country and year. The total value is obtained by adding up the country estimates.

Household survey data on child survival

The Multiple Indicator Cluster Surveys (MICS) are supported by UNICEF and carried out by national governments. The current round of MICS surveys was implemented in over 55 countries in 2005-06. These surveys, along with the USAID-supported Demographic and Health surveys (DHS), are the largest single source of household survey based child mortality estimates and child survival intervention coverage estimates.

Global results

The most recent estimates of under-five mortality, based on the work of the Inter-agency Child Mortality Estimation Group, indicate that the global number of children dying before age five has reached a record low, falling below 10 million per year to 9.7 million in 2006, down from around nearly 13 million in 1990.

Of the 9.7 million children who perish each year, 4.8 million are from sub-Saharan Africa and 3.1 million are from South Asia, (see graph below). The highest rates of under-five mortality are still found in West and Central African countries with 186 deaths per 1,000 live births. In Southern Africa hard won gains in child survival have been undermined by the spread of HIV and AIDS. By contrast, the Latin American and Caribbean region (27 deaths per 1,000 live births), Central and Eastern Europe and the Commonwealth of Independent States (27 per 1,000), and East Asia and the Pacific (29 per 1,000) are on track to reach the Millennium Development Goal of a two-thirds reduction in under-five mortality between 1990 and 2015. In industrialized countries there are just six deaths per 1,000 live births (see graph and table below).

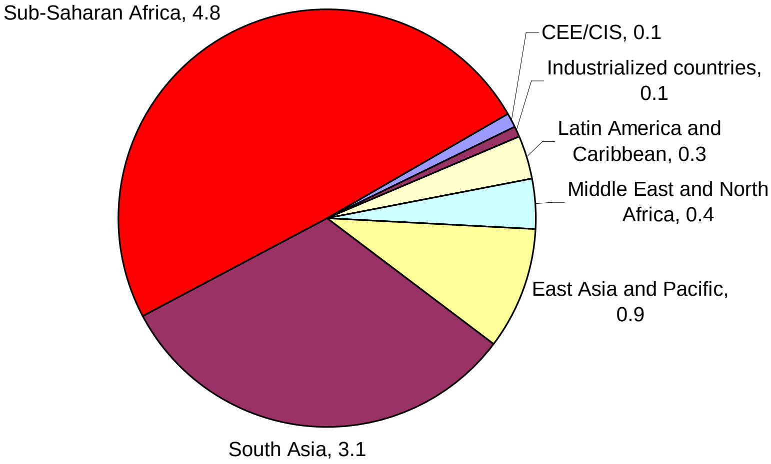
Recent data (1995-2006) collected via household surveys, such as the UNICEF-supported Multiple Indicator Cluster Surveys (MICS) and the USAID-supported Demographic and Health Surveys (DHS) also indicate that under-five mortality is considerably higher among children living in rural areas and for those living in the poorest households.

---

[1] Hill Kenneth, R. Pande, M. Mahy, G. Jones, 1999. Trends in Child Mortality in the Developing World: 1960-1990. UNICEF. New York, NY 10017 USA.

In addition, the latest round of MICS surveys indicate that solid progress has been made in increasing coverage for a range of child survival interventions, such as early and exclusive breast feeding, immunization, Vitamin A supplementation and the use of insecticide-treated bed nets to prevent malaria. The decline in child mortality is in line with earlier reports of progress in measles mortality, which fell by 60 per cent worldwide between 1999 and 2005, and by 75 per cent in sub-Saharan Africa.

The fall in under-five mortality is just one of many positive trends that will be presented in forthcoming publications such as *Progress for Children*, which will chart progress made since the UN General Assembly Special Session on Children in 2002, and *The State of the World's Children* Report which will outline the need for integrated, community-based health services to accelerate progress on child survival.

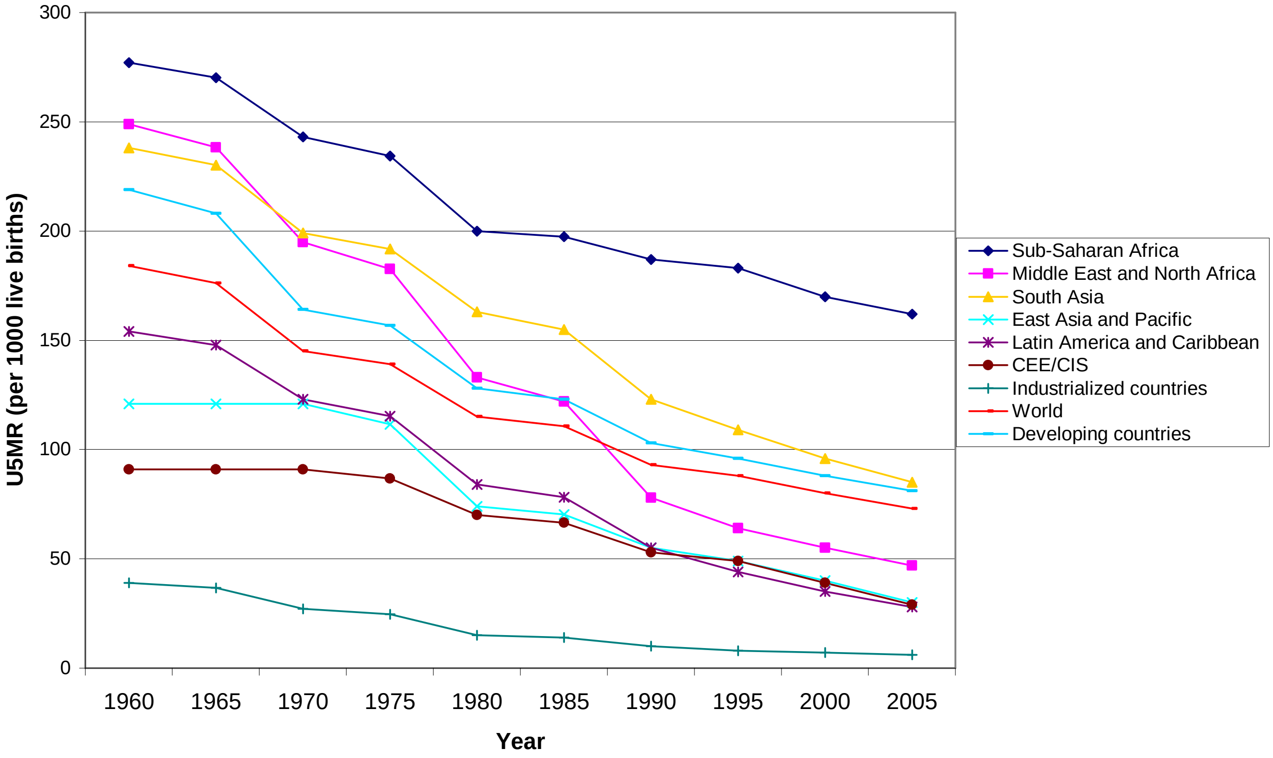


9.7 million children die in 2006 before their fifth birthday

Estimated number of deaths (in millions) occurring before the age of 5 in 2006

Under-five mortality rates (U5MR) have declined during the period 1960-2005

Trends in the probability of dying before the age 5 by regions, 1960-2005



	Under 5 mortality rate		
	1970	**1990**	**2006**
UNICEF's Regions			
Sub-Saharan Africa	243	187	160
Eastern and Southern Africa	220	165	131
West and Central Africa	264	208	186
Middle East and North Africa	195	78	46
South Asia	199	123	83
East Asia and Pacific	121	55	29
Latin America and Caribbean	123	55	27
CEE/CIS*	91	53	27
Industrialized countries	27	10	6
Developing countries	164	103	79
Least developed countries	244	180	142
World	145	93	72

	Under 5 deaths [in millions]		
	1970	**1990**	**2006**
Sub-Saharan Africa	3.2	4.1	4.8
Eastern and Southern Africa	1.5	1.8	1.9
West and Central Africa	1.7	2.4	2.9
Middle East and North Africa	1.3	0.8	0.4
South Asia	5.6	4.7	3.1
East Asia and Pacific	4.9	2.0	0.9
Latin America and Caribbean	1.3	0.6	0.3
CEE/CIS*	0.6	0.4	0.1
Industrialized countries	0.4	0.1	0.1
Developing countries	16.7	12.5	9.6
Least developed countries	3.6	4.0	4.1
World	17.3	12.7	9.7

* Central and Eastern Europe and the Commonwealth of Independent States

ANNEX I

DATA SOURCES AND ESTIMATION METHODS

In countries with accurate registers of births and deaths, infant mortality year by year is measured as the ratio of deaths under one year to births in the same year obtained from civil registration data. The mortality of children after infancy is typically obtained from civil registration information on deaths of young children by age, and population census information on the size of the population of those ages exposed to the risk of dying[1]. Thus civil registration data provide all the information needed to measure infant mortality, which can therefore be readily calculated annually, but measurement of mortality after infancy requires additional information on population sizes.

In countries where the registration of vital events is not complete, the registration of deaths is often less complete than the registration of births, with the result that the registered infant mortality rate underestimates the true value. In these countries, estimates of infant and under-five mortality are typically obtained instead from one or more of three types of survey data. Most similar to registration data is the longitudinal or prospective sample survey. A sample of the national population is followed over a period of time, with all vital events being recorded. Such data provide the basis for calculating the conventional infant mortality rate as the ratio of infant deaths to births, and also provide the basis for calculating mortality rates after infancy since population numbers are also available. Such surveys have not been widely used, partly because they are expensive to mount, and partly because they require careful supervision over an extended period to provide good data.

The second approach is to carry out a sample survey that collects birth histories, with a mother being asked for information on the date of birth and, if relevant, the age at death of every live-born child she has had. Such data have been widely collected by the World Fertility Survey programme in the 1970s and early 1980s, and more recently by the Demographic and Health Surveys project. Both infant and under-five mortality rates can be calculated from the data, dividing deaths for given ages and time periods by exposure to risk in terms of person-years of life lived by the reported children. However, the collection of such information by surveys is complex and requires high levels of interviewer quality and training. The surveys are therefore quite expensive and can only cover small samples.

The third approach is to use what are often called "Brass methods", after William Brass who developed the methodology. Each woman surveyed is asked for very simple information: her age, the total number of children she has borne, and the number of those children that have died. For a particular age group of women, the proportion of children dead depends primarily on two things: the level of under-five mortality, and the distribution of the children by how long they have been exposed to risk. The Brass method (and developments of it) adjust the proportions dead by age group of mother for an estimated exposure distribution in order to arrive at pure measures of under-five mortality and of reference dates for these measures. The adjustment process assumes certain patterns of fertility and under-five mortality by age, and results can be quite sensitive to the choices made.

A more complete review of methods of measuring under-five mortality in countries lacking accurate vital registration data can be found in chapter 1 of *Child Mortality Since the 1960s*, and a review of their various advantages and disadvantages can be found in Hill (1991).

---

[1] Theoretically, in a population with perfect registration of vital events and no migration, the mortality of children after infancy could also be measured purely from civil registration data. However, in practice, registration is not perfect, and migration is non-negligible, so mortality in childhood is calculated from deaths from civil registration data and exposure to risk from a census or population estimate.

A. DATA USED

The basic data have been extracted from *Child Mortality Since the 1960s* in the form of estimates of the infant mortality and under-five mortality rates and the dates to which these estimates refer. For 72 countries, more recent data from new DHS surveys or from newly-available census, civil registration or survey data have been added to update *Child Mortality Since the 1960s.* For these 72 countries, the new mortality estimates are shown in Appendix I in the same format as that used in *Child Mortality Since the 1960s.*

A total of 187 new sets of estimates (DHS direct and indirect estimates, censuses, updated vital registration estimates, etc.) not listed in *Child Mortality Since the 1960s* are used in this volume to estimate trends in IMR and U5MR from 1960 to 1996; of these, 28 sets of estimates were used in Hill and Yazbeck (1994). This volume also includes trends for 12 countries not listed in *Child Mortality Since the 1960s* or in Hill and Yazbeck (1994), for which data were not available earlier, namely Chad, Eritrea, Ethiopia, Gambia, Guinea, Kazakhstan, Laos, Mauritania, Mongolia, Oman, Swaziland, and Yemen.

B. METHODOLOGY FOR ESTIMATION OF TRENDS

There are many ways in which a set of estimates can be obtained from a series of observations, and in which extrapolations forward or backward to any desired time point can be made. The simplest procedure is hand smoothing: drawing a freehand curve through a set of observations, and extending its general trend onwards to some specified time point for which an estimate or projection is required. Such a procedure is unlikely to be objective -- different analysts would almost inevitably draw different lines, particularly for extrapolations beyond the earliest or latest observations. An alternative would be to use hand smoothing for the period covered by observations, and then apply a model to extrapolate; this is essentially what was done in both the United Nations, 1983 and United Nations, 1988 reports on under-five mortality in the developing world. Using this approach, the hand smoothing will still be subjective, and it will be hard to specify the extrapolation model satisfactorily.

Regression analysis offers a set of possible approaches: robust regression, locally-weighted least squares, weighted least squares, or ordinary least squares. Such regression techniques offer a greater degree of objectivity than hand smoothing, but still require the choice of model specification. Each approach has advantages and disadvantages. Ordinary least squares makes few assumptions about the nature of errors, but the results are heavily influenced by outliers. Robust regression essentially underweights observations that have an inordinate influence on the fitted line, but no prior information about the likely accuracy of particular types of observation is used. Locally-weighted least squares fits a non-parametric trend to a set of observations, but provides no basis for extrapolation. Weighted least squares requires the use of pre-determined weights, introducing a subjective element into the analysis, but at the same time allowing the incorporation of prior knowledge about the likely accuracy of different measurement approaches.

The approach adopted here is to fit a regression line to the relationship between infant or under-five mortality rates and their reference dates using weighted least squares. The basic model assumes that the *rate* at which infant or under-five mortality changes is linear in time, that is, that mortality risk changes at a constant annual percentage rate over some defined time period. The dependent variable is the logarithm of either the infant mortality rate or the under-five mortality rate. The independent variables used are date variables. The simplest model simply relates the logarithm of each estimate of the risk of dying to the date of the estimates; this model implies a constant rate of change in mortality over the entire period studied. Figure 2(a) shows a hypothetical example of such a simple model for under-five mortality. More complex models can allow the rate of change of mortality to vary over the period; Figure 2(b) shows an example where the relationship between the under-five mortality rate and time is allowed to change three times, reflecting three "knot" points (defined every time that a sum of weights reaches five, as explained below). Clear changes in trend are indicated in this instance.

The model used in this volume allows the rate of change of infant or under-five mortality to vary according to the number of independent observations available. This model, shown below as equation (1), includes an underlying date variable and additional variables measuring time since a series of "knot" dates. The rate of change of infant or under-five mortality can change at each knot. The number and location of these knot dates are determined by the number and location in time of observations available for a particular country. Knots are defined working backwards in time from the most recent observation. The weights for successive observations are summed, and a knot is defined every time that the sum of weights reaches a multiple of five. This definition of a knot is based on the idea that five years of vital registration, or one set of indirect or direct estimates such as one DHS, are each sufficient to define one trend or slope. The weights assigned to five years of vital registration, and to one set of direct or indirect estimates, each sum to five. Thus, the process of allowing a knot to be defined when weights sum to five, means, in effect, that each DHS, WFS or other such survey, and each five-year period of vital registration defines a particular slope[2].

The only exception to this procedure is for the last knot defined (i.e., the earliest knot in time). For this last knot, the remaining weights must sum to at least five, to ensure that the first date variable is based on observations whose weights sum to at least five. The more acceptable observations there are for a particular country, the more knots there will be, and the more flexible the trend in infant or under-five mortality over time will be.

The underlying model used is

$$\ln({}_{n}q_{0})_{i} = b_{0} + b_{1}(\mathit{date}) + b_{2}(\mathit{postk1}) + b_{3}(\mathit{postk2}) + b_{4}(\mathit{postk3}) + \ldots + e_{i} \qquad (1)$$

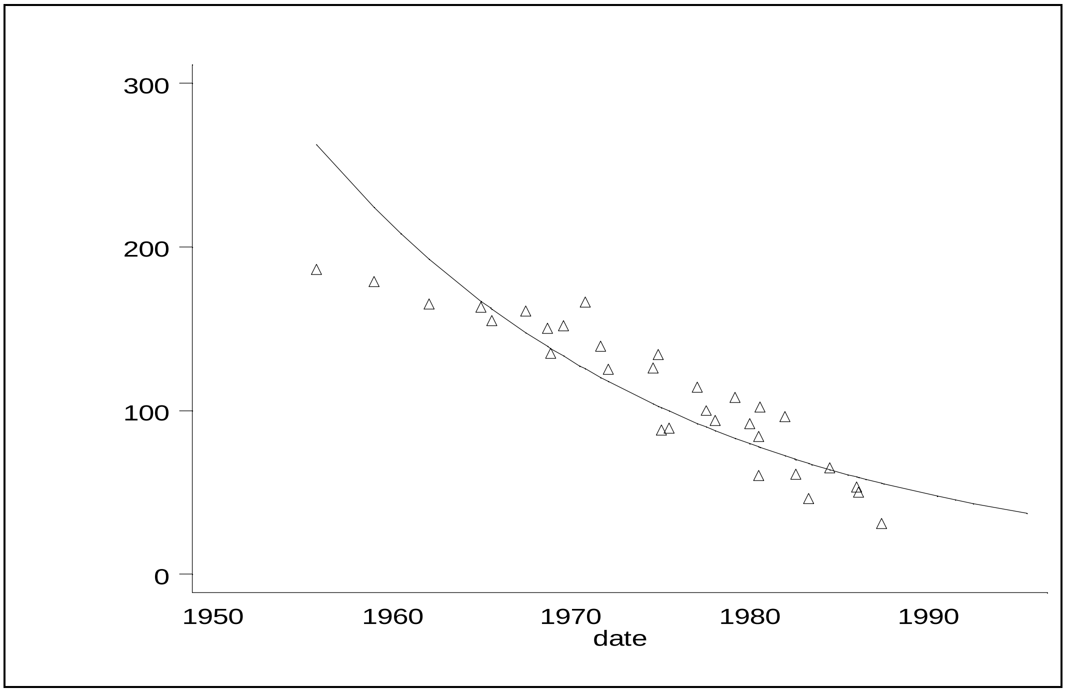
The variable *date* is simply calendar year; *postk1* is date minus the date of the earliest defined knot if positive, or zero otherwise, and picks up any change in trend after the first knot (note that the knots are

---

[2] See following section, *Weights*, for details on weights assigned to particular sources and kinds of data estimates.

Figure 2: Fitted Trends Using Different Numbers of Knot Parameters

a) One date and no knot parameters



b) One date and three knot parameters

defined from the present backwards into the past, but the earliest knot is defined to ensure at least five observations between it and the start of the series); *postk2* is date minus the date of the second defined knot if positive, or zero otherwise, and picks up any change in trend after the second knot; and so on. Thus, the number of slope-changing time variables varies with the number and weight of the observations over time. The coefficients on *postk1, postk2,* etc. can be interpreted as changes in the rate of change of infant or under-five mortality with time in that particular period. Thus the rate of change in period 1 is b_1; in period 2, $(b_1 + b_2)$; in period 3, $(b_1 + b_2 + b_3)$; and so on. Using the examples in Figure 1, only a single date variable with no slope-changing variables would be used to fit a trend in U5MR for the Congo, since the data weights only sum to five. For Colombia, the date variable and seven slope-changing variables would be used, since the data weights sum to 39.8.

It should be noted that the error term e_i is assumed to be normally distributed around the logarithm of the mortality indicator. As a result, estimates of mortality obtained by exponentiating an estimated value of the logarithm of the mortality indicator will be biased upwards by an amount that will depend on the goodness of fit of the model. This is a relatively benign bias in the sense that the infant or under-five mortality estimates obtained will be conservative, and the poorer the fit of the model the more conservative the estimates will be.

Weights

Weighted least squares regression is used to fit equation (1) to each country's data. Weighted least squares is used because a substantial body of evidence suggests different validity weights for different types of observations. For example, it is generally thought that the quality of retrospectively reported information deteriorates with the length of time ago of the events reported (Som, 1973).

Each estimate from vital registration or a prospective survey is given a "standard" validity weight of 1.0. For vital registration, the weight is justified by the typically large number of events involved and by the lack of any substantial lag between event and report; for prospective surveys, the weight is justified by the lack of lag and by the accuracy enforced by the data collection methodology. Estimates derived from birth histories are assigned standard weights that vary with the length of time before the survey to which the estimate refers, on the grounds that recent information is more likely to be accurate than information for periods further in the past. Specifically, estimates for the five years before the survey are given a weight of 2.0, for periods five to nine years before the survey, 1.8, and periods 10 to 14 years before the survey, 1.2.

Weights for indirect estimates based on the proportions dead of children ever born vary by age group of mother; estimates based on reports of young women are given low weight, zero for women age 15 to 19, and 0.2 for women ages 20 to 24, because of the well-known selection problems that affect such estimates (early childbearing is often highest among the poor, who also suffer the highest under-five mortality rates). Estimates based on reports of women aged 25 to 29, 30 to 34 and 35 to 39 are given the highest weights, 1.2 each. Then, as age increases the weights decline slowly, on the grounds that information about events longer ago is more prone to error, such that estimates for the age group 40 to 44 get a weight of 0.8, and those for the age group 45 to 49 get a weight of 0.4.

Note that it is these weights that are summed to determine the location of knots described in the preceding section. The under-five mortality estimates from a particular survey collecting birth histories have a combined weight of five, as do the estimates from a particular survey providing indirect measures. In cases where a particular survey contributes two types of estimates (for example, a DHS survey that provides both direct estimates from the birth history and indirect estimates from children ever born/children surviving) the two sets of weights are reduced to half their standard level, so that the data source is not overly weighted.

The observation-specific weights described above are essentially based on the authors' judgment and experience. However, robust regression techniques can be used to estimate robust weights for particular types of observations; these estimated robust weights can then be compared with the observation-specific weights given above. The robust regression technique incorporated in the statistical package STATA has been used here. This technique first estimates an unweighted regression and excludes observations identified as gross outliers. The routine then works iteratively with the remaining observations, assigning weights on the basis of absolute residuals from the regression line (StataCorp, 1995.)

Taking into account geographic representation, a total of 13 countries with a large number of different types of observations -- particularly indirect estimates based on children ever born and children surviving and direct estimates based on birth histories -- were selected for this exercise. For each country, a model of the type of equation (1) was robustly fitted using the same independent date variables as used in the standard analysis. The robust weights obtained for each type of point -- for example, an indirect estimate based on reports of women age 15 to 19, or a direct estimate based on events in the period five to nine years before the survey -- were recorded. Table 2 shows the first quartile, median, and third quartile values of the weights, and the interquartile range as an indicator of reliability of a particular type of point.

The median robust regression weights for indirect estimates are higher than those used in the standard analysis, though they show a broadly similar pattern with age. The interquartile range varies from very high for reports of women ages 15 to 19 to a mere 0.06 for women ages 30-34 and 35-39, confirming the unreliability of estimates based on reports of women aged 15-19, and the higher reliability of estimates based on reports of women in their thirties.

One surprising feature of the indirect weights is how well the estimates based on reports of older women hold up: with medians no lower than those for observations based on reports of women age 30-34 or 35-39, and interquartile ranges only modestly higher, there seems on this basis to be very little quality loss in the reports of older women. The weights for direct estimates are also somewhat higher than those used in the regular analysis for periods other than the five years before the survey, and decline only slightly, on the basis of their median, for periods longer before the survey. However, the interquartile range increases substantially with time ago, providing some support for the standard weights. On the basis of the medians, direct and indirect estimates seem to be about equally satisfactory, though the interquartile ranges for the direct estimates tend to be higher than those for indirect estimates based on reports of women aged 25 or more.

Table 2: Weights of Certain Types of Data Points as Estimated by Robust Regression[1]

Type of Observation	Weight for: First Quartile	Median	Third Quartile	Interquartile Range
Indirect Estimates Based on Age Groups:				
15-19 (81)	0.00	0.25	0.89	0.89
20-24 (88)	0.68	0.90	0.96	0.28
25-29 (89)	0.90	0.97	1.00	0.10
30-34 (89)	0.95	0.98	1.00	0.05
35-39 (89)	0.95	0.98	1.00	0.05
40-44 (89)	0.93	0.98	0.99	0.06
45-49 (87)	0.89	0.97	0.99	0.10
Direct Estimates Based on Years Before Survey:				
0-4 (31)	0.88	0.97	1.00	0.12
5-9 (31)	0.82	0.95	0.99	0.17
10-14 (31)	0.75	0.94	0.97	0.22

1. The countries included in this analysis were as follows:
Latin America: Brazil, Colombia, Costa Rica
Sub-Saharan Africa: Kenya, Senegal, Zimbabwe
Middle East-North Africa: Jordan, Tunisia, Turkey
South Asia: Bangladesh, Sri Lanka
East Asia and Pacific: Indonesia, Thailand
The outcome variable used in the model was the under-five mortality rate in all cases. The numbers in parentheses are the numbers of observations of points of that sort included in the combined data sets.

Applying the methodology

The intention of the methodology is to provide a transparent and largely objective way of fitting a smoothed trend to a set of observations, and of extrapolating the trend to cover the period from 1960 to the present. However, there are subjective judgments involved at every step.

Step 1 of the smoothing and extrapolation process fits equation (1) using appropriate date variables and standard validity weights as defined above. The infant mortality rate and the under-five mortality rate are fitted independently. The observations and fitted line are displayed graphically. Results from step 1 depend on the standard validity weights applied to different types of data. As shown in Table 2, robust regression techniques give some support to the weights used, though the standard weights tend to be lower than those estimated robustly. The standard validity weights are objective in the sense that they are invariant across data sets or across countries, but they are subjective in the sense that different analysts would choose different values.

In step 2, results from step 1 are critically examined and data sets that are clearly aberrant are identified, such as vital registration sequences that fall consistently below all other infant mortality estimates, or indirect estimates that are clearly inconsistent with the majority of other sources. The weights for the entire aberrant data set (except in some cases for the registered infant mortality rate, where the sequence becomes consistent with other estimates at some time point) are reduced by a constant factor that is generally zero (giving no weight to estimates from that data set). Equation (1) is fitted once more using these revised weights.

Figure 3 shows the estimates of under-five mortality for Egypt from step 1 and step 2 curves. In the case of Figure 3a, the trend line for Egypt is estimated using standard weights for all sources. From the graph, it is clear that the estimates from the 1976 and 1986 population censuses are out of line with all other survey estimates, thus pulling the trend line down and giving the (almost certainly erroneous) impression of rising under-five mortality in the 1960s. To avoid this problem, the estimates from the two censuses were all given zero weight in step 2, shown in Figure 3b. Inspection of step 2 results can lead to further modifications of data set weights for a third step, but such further iteration was unusual.

This second step involves subjective choices as well. Most important is the decision as to whether, and if so by how much, to underweight entire data sets. As the example of Egypt in Figure 3 shows, the decision to underweight some data sets relative to others can have a very large effect on the sequence of under-five mortality estimates obtained. In the process followed in this volume, this decision is made on the basis of graphical inspection. If the estimates from one source are clearly higher or lower than the bulk of available estimates or if their time trend is clearly different a constant factor of zero (giving zero weight to the entire data set) might be used. If the estimates fluctuate a lot or display some other undesirable characteristic a constant factor between zero and one might be used, thus reducing the standard weight for the entire data set. It is this decision which is most likely to introduce substantial differences in results between analysts. In the analysis reported here, it is assumed that errors are more likely to result in underestimates of infant or under-five mortality than in overestimates. Thus when two data sets indicate very different levels, the set indicating higher mortality is assumed, other things being equal, to be more likely to be right. In the case of Egypt, the lower estimates from the two censuses were rejected in favor of the higher estimates from the series of birth history surveys.

The final decision that has to be made is whether to select the infant mortality or the under-five mortality sequence of estimates as the more consistent series. The two sequences are estimated independently and there is no reason why they should be consistent through their range of observation or extrapolation. In practice, one of the two sequences is chosen, partly on the basis of the graphs and partly taking into account the nature of the evidence. For example, a sequence based largely on indirect estimates is likely to give stronger estimates of under-five mortality than of infant mortality because of the sensitivity of indirect estimates of infant mortality to the age pattern of childhood mortality used in their derivation. Similarly, a sequence based on birth histories is likely to give better estimates of under-five mortality than of infant mortality because of the common problem in birth histories of rounding reported age at death up to one year, thus reducing infant mortality, but not under-five mortality. On the other hand, a sequence based largely on vital registration of reasonable quality is likely to be stronger for infant mortality.

In the graphs displayed in this volume, the estimates of infant (IMR) and under-five (U5MR) mortality rates from step 2 are examined graphically to ascertain which of the two sets of estimates (IMR or U5MR) is to be adopted as the more consistent one, given the nature of available information. Once one of the two indicators has been selected, corresponding values of the other (the derived indicator) can be obtained through a model life table system[3]. The final graphs display the trend line from step 2 for the selected indicator; the trend line showed

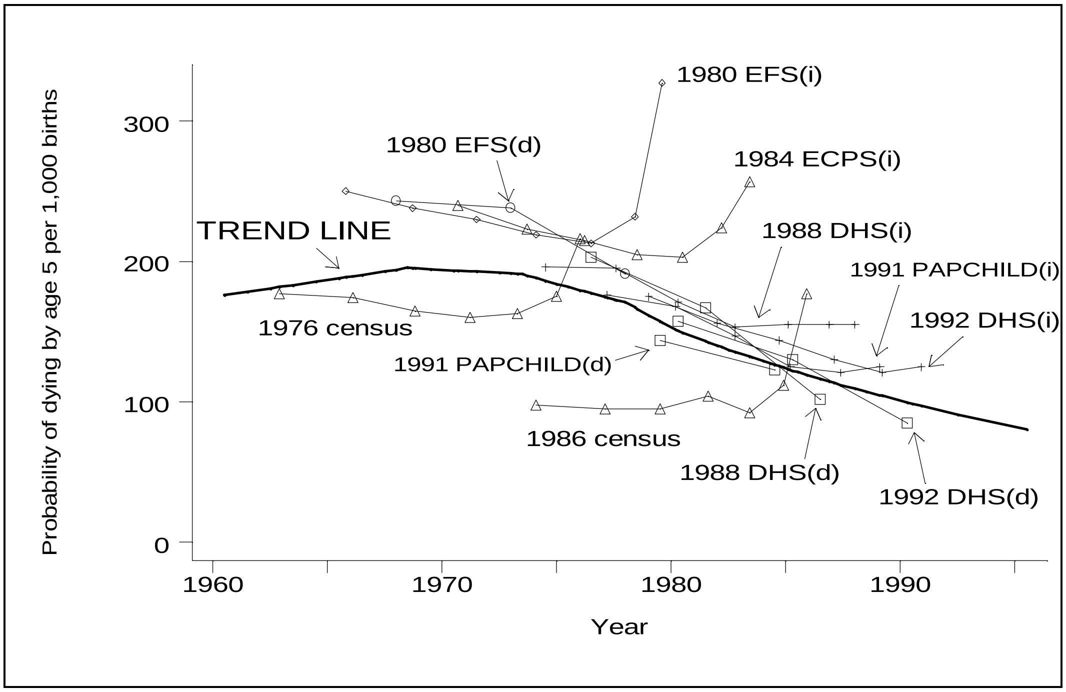
---

[3] The indirect estimates given in *Child Mortality Since the 1960s* use a Coale-Demeny model life table family in almost all cases; the same family as used in *Child Mortality Since the 1960s* is used in this volume to estimate the derived indicator from the selected one. Where the estimates of U5MR fall below 25 per 1000, the West model life table is used.

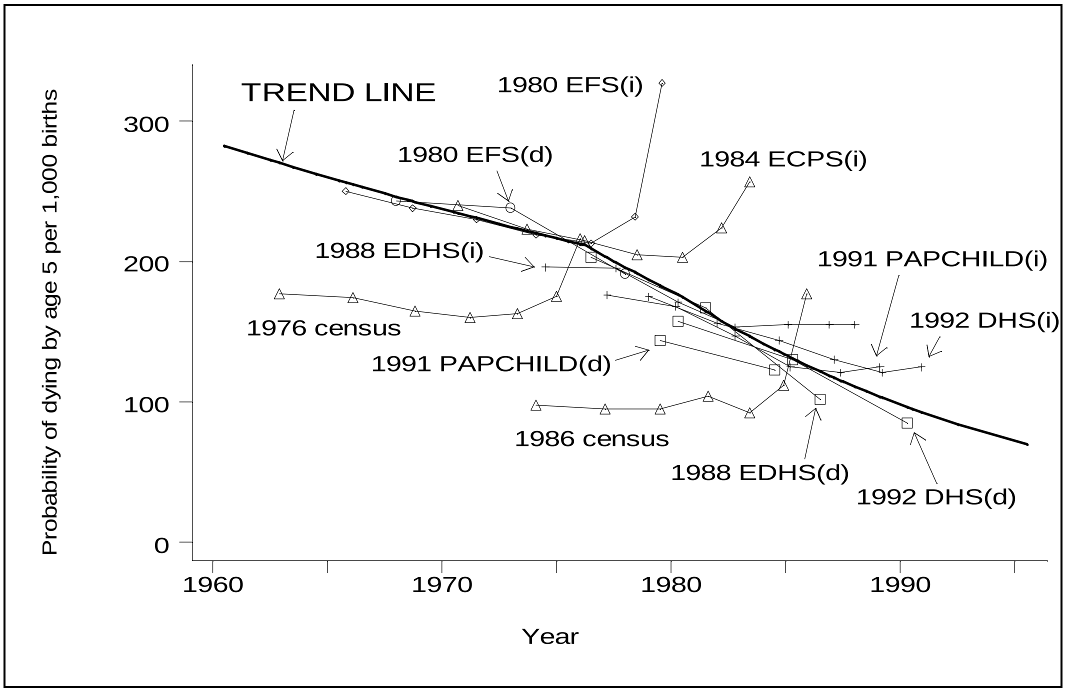
for the derived indicator is based on the values implied by the trend line of the selected indicator. Say, for example, that inspection of the graphs suggests that the U5MR sequence is to be preferred. The trend in infant mortality shown in the country graph will then be based on the IMRs implied by the U5MR trend using the selected model life table family, rather than on the observed IMRs.

Figure 3: Illustration of the Reweighting of Certain Data Sets: Egypt

a) Trend line and data sets using initial weights for the trend



b) Trend line and data sets using zero weights for 1975 and 1986 censuses for the trend



C. METHODOLOGY FOR ESTIMATION OF NUMBERS OF DEATHS IN 1996

Estimating numbers of under-five deaths and numbers of births

The estimates of childhood mortality in this volume are given in the form of infant mortality rates or under-5 mortality rates. To estimate the number of deaths of children under 5 in a given year, a first approximation could be obtained by applying the U5MR to an estimate of the number of births in that year. However, if the population is growing (specifically, if the number of births is increasing each year) this approximation will slightly over-estimate deaths because the mortality risk at each age under 5 is implicitly being applied to a number of children at risk derived from the number of births in the current year. The actual number at risk will be smaller, because the number of, for example, four year olds is the number of survivors of a smaller number of births four years earlier. If the number of births is declining each year, the opposite effect will apply, and the number of deaths obtained by applying the U5MR to the annual births will under-estimate deaths under 5.

Simulations with models were used to explore the relationship between the number of births in a given year, the growth rate of the number of births over the preceding five years, the U5MR and the number of under five deaths, U5D. These simulations used the Coale-Demeny (1983) model life tables (survivors to exact ages 1, 2, 3, 4 and 5) for each family and mortality level in combination with a range of growth rates of the birth stream (0, 1, 2 and 3 percent) to estimate the true number of U5D for a given year. For any of these simulated populations, the U5D is related to the number of births by a factor k_t.

$$U5D_t = (B_t * U5MR_t) * k_t$$

where $U5D_t$ is the under-five deaths in year t, B_t is the births in year t (in '000s), and $U5MR_t$ is the under-five mortality rate per 1,000 for 1996 estimated in this volume.

The factor k_t will depend on the age pattern of mortality and the rate of growth of births, r_b.

$$k_t = (a_f + b_f * r_b)/100{,}000.$$

The underlying pattern of child mortality is reflected by a_f and b_f, which are constants estimated by regression analysis from the simulations mentioned above, for a given family of Coale-Demeny model life tables for a given percent growth rate of births, r_b. Values of a_f and b_f estimated by regression are shown by Coale-Demeny model family in Table 3. It may be mentioned in passing that the regression of k_t on U5MR and (U5MR * r) explains essentially all the variability of k_t in the models used (though it must be remembered that the models assume a fixed rate of growth of births over a five year period preceding the year of observation, so real populations will have more variability than the simulations).

Table 3: Constants for Use in Converting Births and U5MR's into Numbers of Deaths Under Five

	Coale-Demeny Model Family			
Parameter	*North*	*South*	*East*	*West*
a_f	0.99985	0.99987	0.99992	0.99989
b_f	-0.01252	-0.01193	-0.00918	-0.01043

Numbers of births in year t can be estimated from standard United Nations population projections as an average of the average annual number of births in the five years preceding t and the average annual number of births in the five years following t. Thus

$$B_t = B_{t-5,t} * \exp(2.5 * r_b(t-5,t))$$

where

$$r_b(t-5,t) = 0.1*(\log_e(B(t,t+5)/B(t-10,t-5)))$$

where the U5MR is expressed per 1,000 live births and the growth rate of births r is expressed in percentage form.

The United Nations' publication *World Population Prospects 1994* (1995) gives national population estimates and projections for calendar years ending in zero and five. Annual numbers of births are given for the five year intervals between such calendar years[4]. The growth rate of births is unlikely to change fast enough to justify fine tuning, so for years that are not multiples of five, it is suggested that the growth rate of births for the five year periods closest to the one required be used. Thus, for example, for years 1993 to 1997, the growth rate used should be the average of the annual growth rate from 1985-90 to 1990-95 and the annual growth rate from 1990-95 to 1995-2000.

In summary, the procedure used here to estimate the number of deaths in 1996 (and the process for other years would be similar) is to use information from the United Nations estimates and projections, country by country, to estimate the number of births in 1996 for each country, and combine these country-specific estimates of births with the estimates of U5MR developed in this volume, adjusting for any changes in the rate of growth of births in the years preceding 1996. The adjustment factor, k_t, is close to 1, being slightly less than 1 when births are increasing and slightly greater than 1 when births are decreasing. The calculations are illustrated below in a worked example.

---

[4] However, for the worked example that follows, and in Appendix I, we used data on births in 1985-90, 1990-95 and 1995-2000 provided by the UN Population Division to UNICEF; these data differ only slightly from the *World Population Prospects* (1995) estimates.

A worked example

Taking the example of Bangladesh, data on births provided by the UN Population Division to UNICEF give the estimated annual births 1985-90 as 3,825 thousand, and the projected numbers for 1990-95 and 1995-2000 (using the medium variant of the projections) as 4,051 thousand and 4,190 thousand respectively. The growth rate for the period 1990-95 can then be approximated as the average of the growth rate for 1985-90 to 1990-95 and that for 1990-95 and 1995-2000:

$$r_b(1990\text{-}95) = 0.5*0.2*(\log_e(\text{Births}_{1990\text{-}95}/\text{Births}_{1985\text{-}90}) + \log_e(\text{Births}_{1995\text{-}2000}/\text{Births}_{1990\text{-}95}))$$

$$= 0.1*(\log_e(\text{Births}_{1995\text{-}2000}/\text{Births}_{1985\text{-}90}))$$

$$= 0.1*(\log_e(4190/3825))$$

$$= 0.00911 \text{ (or 0.911 per cent)}$$

The estimated U5MR for Bangladesh in 1995 is 122, and the Coale-Demeny family used is `South', so k is estimated as

$$k_{1995} = 0.99987 - 0.01193*0.911 = 0.98900$$

The 1995 births are estimated by applying the 1990-95 to 1995-2000 growth rate of births to the 1990-95 annual number of births:

$$B_{1995} = \text{Births}_{1990\text{-}95} * \exp(2.5 * r_b(1990\text{-}95))$$

$$= 4051*\exp(2.5*0.00911)$$

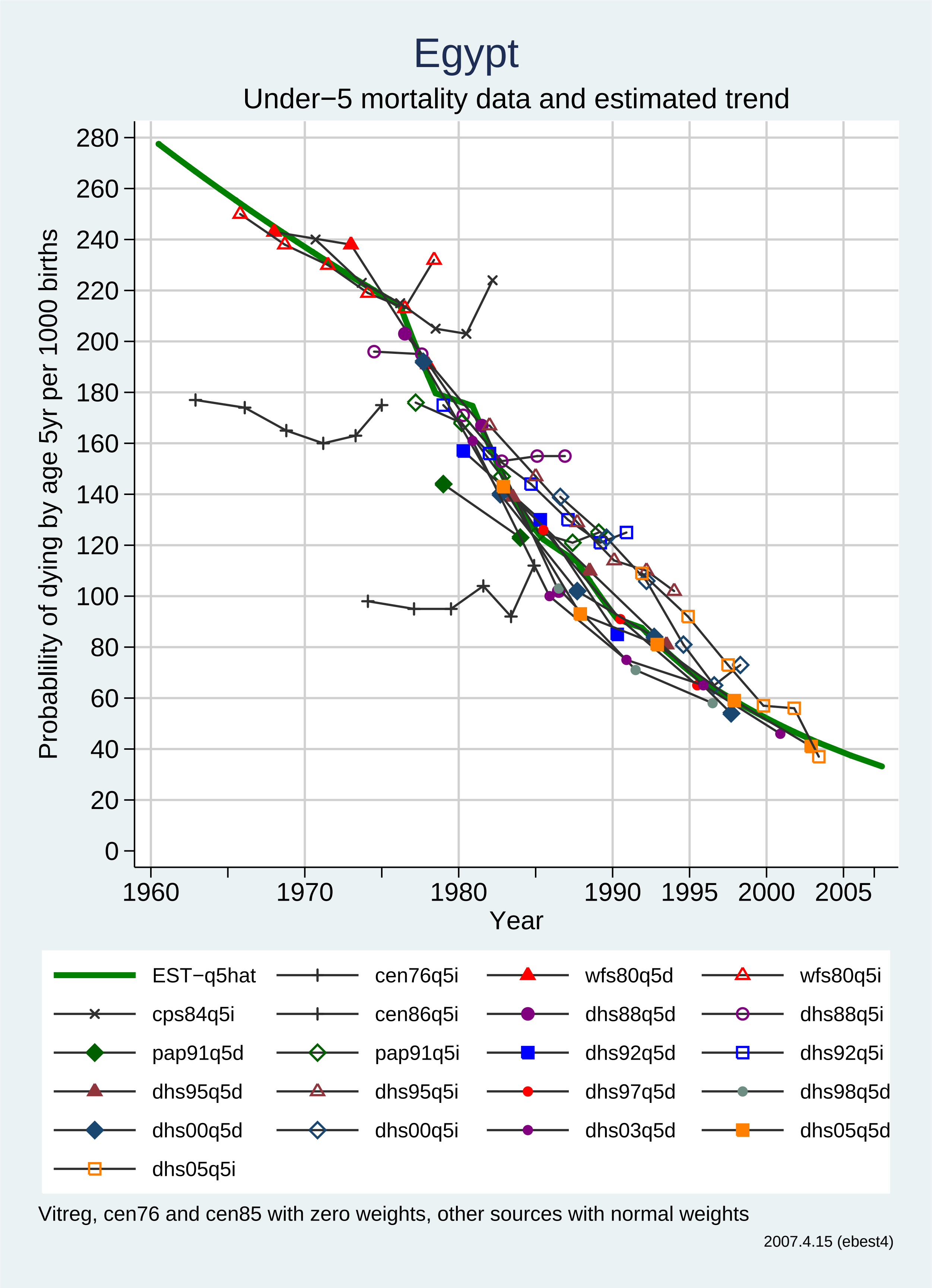
$$= 4144.3$$

so

$$U5D_{1995} = 4144.3 * 115 * 0.98900 = 471{,}352$$

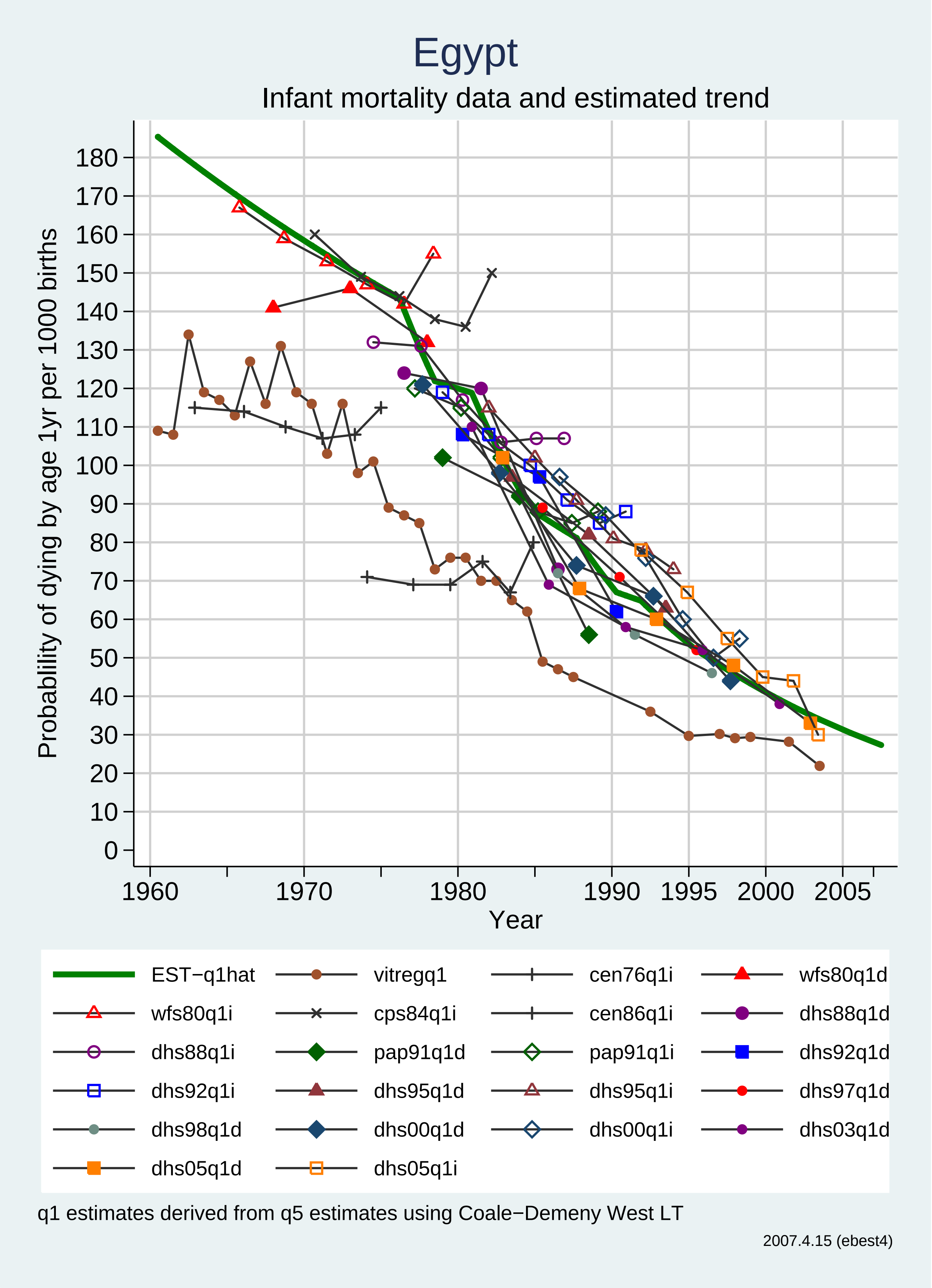
Note that the product of the U5MR and the estimated 1995 births would have given a slightly higher total, 476.6 thousand.

ANNEX II



Vitreg, cen76 and cen85 with zero weights, other sources with normal weights

2007.4.15 (ebest4)



Egypt

Infant mortality data and estimated trend

q1 estimates derived from q5 estimates using Coale−Demeny West LT

2007.4.15 (ebest4)

Egypt child mortality estimates

Final version - 2007.4.15

date	q1hat	q5hat	q1reg
1960.5	185.4	277.5	169.3
1965.5	170.5	255.4	159.4
1970.5	157.2	235.1	150.0
1975.5	145.1	216.5	141.2
1980.5	119.3	175.5	117.7
1985.5	86.6	122.3	84.3
1990.5	66.7	90.6	67.2
1995.5	52.0	68.1	53.4
2000.5	40.0	50.6	40.6
2005.5	30.5	37.4	30.7
2007.5	27.3	33.2	27.5
	West LT		

Note:
q1hat and q5hat are final estimates for q1 and q5 respectively; q1reg are regression estimates before final step.

Egypt data sources

Acronym	Description
vitregq1	Vital registration IMR data
cen76q1i	Census 1976 indirect IMR data
cen76q5i	Census 1976 indirect U5MR data
wfs80q1d	World Fertility Survey 1980 direct IMR data
wfs80q5d	World Fertility Survey 1980 direct U5MR data
wfs80q1i	World Fertility Survey 1980 indirect IMR data
wfs80q5i	World Fertility Survey 1980 indirect U5MR data
cps84q1i	Contraceptive Prevalence Survey 1984 indirect IMR data
cps84q5i	Contraceptive Prevalence Survey 1984 indirect U5MR data
cen86q1i	Census 1986 indirect IMR data
cen86q5i	Census 1986 indirect U5MR data
dhs88q1d	DHS 1988 direct IMR data
dhs88q5d	DHS 1988 direct U5MR data
dhs88q1i	DHS 1988 indirect IMR data
dhs88q5i	DHS 1988 indirect U5MR data
pap91q1d	PAPCHILD 1991 direct IMR data
pap91q5d	PAPCHILD 1991 direct U5MR data
pap91q1i	PAPCHILD 1991 indirect IMR data
pap91q5i	PAPCHILD 1991 indirect U5MR data
dhs92q1d	DHS 1992 direct IMR data
dhs92q5d	DHS 1992 direct U5MR data
dhs92q1i	DHS 1992 indirect IMR data
dhs92q5i	DHS 1992 indirect U5MR data
dhs95q1d	DHS 1995 direct IMR data
dhs95q5d	DHS 1995 direct U5MR data
dhs95q1i	DHS 1995 indirect IMR data
dhs95q5i	DHS 1995 indirect U5MR data
dhs97q1d	DHS 1997 direct IMR data
dhs97q5d	DHS 1997 direct U5MR data
dhs98q1d	DHS 1998 direct IMR data
dhs98q5d	DHS 1998 direct U5MR data
dhs00q1d	DHS 2000 direct IMR data
dhs00q5d	DHS 2000 direct U5MR data
dhs00q1i	DHS 2000 indirect IMR data
dhs00q5i	DHS 2000 indirect U5MR data
dhs03q1d	DHS 2003 direct IMR data
dhs03q5d	DHS 2003 direct U5MR data
dhs05q1d	DHS 2005 direct IMR data
dhs05q5d	DHS 2005 direct U5MR data
dhs05q1i	DHS 2005 indirect IMR data
dhs05q5i	DHS 2005 indirect U5MR data
IMR	Infant mortality rate
U5MR	Under-5 mortality rate
DHS	Demographic and Health Survey