

**Modified Logit Life Table System:  
Principles, Empirical Validation and Application**

**GPE Discussion Paper No.39**

**CJL Murray  
BD Ferguson  
AD Lopez  
M Guillot  
JA Salomon  
O Ahmad**

**Evidence and Information for Policy (EIP)**

**World Health Organization**

## Introduction

Model life table systems [1] [2] [3] are extensively used in demographic, epidemiological and economic analyses. Probably the most widespread use is to infer age patterns of adult mortality, about which comparatively little is known in developing countries, from levels of child mortality, which are much more reliably documented [4]. Yet, substantial evidence has accumulated that these widely used model life table systems do not adequately represent the range of age-specific patterns that are empirically observed. The routine use of split-level modifications of the Coale-Demeny and the United Nations model life table systems is one manifestation of the inadequacy of the original models for current estimation purposes. Concomitantly, there has been a major expansion of empirically observed data on age-specific mortality in countries with complete or very nearly complete registration systems over the last 30 years [4]. These data provide an opportunity to improve the widely used model life table systems through a reappraisal of age patterns of mortality that have been observed in such populations.

In this paper, we present the development and testing of a new model life table system based on a modification of the Brass Logit life table system [2]. The paper is divided into six sections. The first section briefly reviews some of the main uses of model life tables and consequently the requirements for a good model life table system. Section two reviews the main two-parameter model life table systems, emphasizing the Coale-Demeny, United Nations and Brass systems. In the third section, we present the logic and mathematical foundation for a modification of the Brass Logit life table system. The WHO dataset of high quality life tables which provides the empirical basis for the development of the WHO system is reviewed in Section 4. The empirical estimation of the Modified Logit life table system is developed in section 5, including basic information on the robustness of the model. Section 6 provides a direct empirical assessment of the adequacy and predictive power of the Coale-Demeny and Modified Logit systems for a random sample of life tables. Some limitations and implications of this work are presented in the final section.

## I. Uses and Required Properties for Model Life Table Systems

Understanding the strengths and weaknesses of model life table systems and thus the direction for an improved system should start with a clear articulation of the multiple uses of model life tables. Model life tables are extensively used for smoothing data, incorporating age-specific mortality patterns in various indirect estimation techniques such as sibling or parental survival, and forecasting age-specific mortality rates [1] [3] [5].

One of the most important uses of model life tables is in routine demographic estimation work in settings where complete vital registration is not available. A complete life table often is estimated with information only on child mortality or child mortality and some measure of adult mortality experience. Another important use of model life tables is in the economic appraisal of health interventions when the benefits of an intervention must be modelled in the context of general levels of mortality.

Model life tables are not models in the usual sense of the word. They are not causal theories or statistical models. Rather, model life tables can be thought of as representation theorems. The central thesis is that the complex phenomenon of age-

specific mortality rates can be adequately represented by two or three parameters such as model family and level. Being able to represent a full schedule of mortality by age with two or three pieces of information simplifies understanding mortality patterns and has proven to have multiple analytical uses in many fields. Thinking of model life tables as representation theorems may help in formulating appropriate empirical tests of the adequacy of a model life table system.

We propose at least three required properties for a model life table system. The first required property is that it be simple and easily used. In practice, this probably means that a model life table system should at most use two parameters to define a unique life table. More complicated systems may do better on the second and third criteria described below but the fact remains that such systems have not been widely used in applied work. We include in the category of two parameter systems: the Coale-Demeny family of life tables, the United Nations models, the Brass Logit system and the Ledermann system [6]. The Coale and Demeny and United Nations systems are *de facto* two parameter systems, the choice of family is one parameter and the level is the second parameter. The Brass logit system when a single global standard is used has two parameters,  $\alpha$  and  $\beta$ . When multiple standards are used, it becomes a three parameter system.

Second, any two-parameter model life table system should also adequately capture the true range of age-specific mortality patterns seen in real populations. In other words, model life table systems should not under-represent the extent to which mortality by age can vary across populations. For example, if one looks at the scatterplot of child mortality measured using  ${}_5q_0$  and adult mortality measured using  ${}_{45}q_{15}$  in populations with good vital registration data, how much of the diversity of this pattern is captured in the model life table system?

Third, when a model life table system is used to select a life table to represent mortality by age for a population, how close a fit is there between the predicted mortality rates and actual mortality rates? The fit between predicted and actual can be assessed in many ways such as the root mean square error in the death rates (or log of death rates), the explained variance or the average relative error in age-specific death rates. Formal assessment of the predictive power of a model life table system should be an absolute requirement to judge its adequacy.

There are other uses and therefore other criteria that can be proposed to evaluate a model life table system. For this paper, however, we focus on two-parameter systems and more formally assess the range of age-specific mortality patterns they capture and their predictive power.

## II. Two Parameter Model Life Table Systems

The basic objective in the creation of any model life table is to construct a system that gives mortality rates by sex and age, defined by a small number of parameters that capture the level as well as the age pattern of mortality. If a particular model adequately represents reality, the characteristics of a given population can be summarized by the parameters of that model, thereby facilitating the study of variation among populations or within a population over time. The principles underlying each of the existing model life tables are discussed below.

The first set of model life tables was published by the United Nations in 1955 [7]. This was a relatively simple one-parameter system indexed on infant mortality levels. Subsequently, the United Nations published a revised set of model life tables in 1981 which attempted to construct regional models, as did Coale and Demeny, but using data from developing countries judged adequate for inclusion in the empirical dataset. The underlying data consisted of 36 life tables covering a wide range of mortality levels from developing countries, by sex. Sixteen pairs of life tables came from 10 countries in Latin America, 19 pairs from 11 countries in Asia, and one pair from Africa. Five families of models were identified, each with a set of tables ranging from a life expectancy of 35 to 75 years for each sex. Each family of models covers a geographical area: *Latin American*, *Chilean*, *South Asian*, *Far Eastern* and a *General Pattern*. The general model was constructed as an average of all the observations.

However, perhaps the most widely used model life table system has been the Coale-Demeny regional model life tables. First published in 1966, they were derived from a set of 326 life tables, by sex, from actual populations. This set included life tables from several time periods (23 from before 1870 and 114 from after the Second World War) and mostly from Western countries. Europe, North America and Oceania contributed a total of 246 tables. Three were from 32 from Asia, 33 from Latin America and 15 from Africa. All of the 326 selected life tables were derived from registration data, and were subjected to stringent standards of accuracy. In addition to discarding life tables which lacked separate treatment of the sexes or breakdown into five year age groups (with separate age groups for age 0 and age 1-4), an effort was made to avoid repetition by excluding sub-national life tables as well as life tables referring to consecutive intercensal years routinely calculated in some countries. The researchers also eliminated tables covering years in which a major war occurred [1].

Further analysis of the underlying relationships identified four typical age patterns of mortality, determined largely by the geographical location of the population, but also on the basis of their patterns of deviations from previously estimated regression equations. Those patterns were called: North, South, East, and West. Each had a characteristic pattern of child mortality. The *East* model comes mainly from the Eastern European countries, and is characterized by high child mortality in relation to infant mortality. The *North* model is based largely on the Nordic countries, and is characterized by comparatively low infant mortality, high child mortality and low old age mortality beyond age 50. The *South* model is based on life tables from the countries of Southern Europe (Spain, Portugal, and southern Italy), and has a mortality pattern characterized by (a) high child mortality in relation to infant mortality *at high overall mortality*, and (b) low child relative to infant mortality *at low overall mortality*. The *West* model is based on

the residual tables not used in the other regional sets (i.e., countries of Western Europe and most of the non-European populations). It is characterized by a pattern intermediate between North and the East patterns. Because this model is derived from the largest number and broadest variety of cases, it is believed to represent the most general mortality pattern. In this system, any survivorship probability, whether from birth or conditional on having attained a certain age, uniquely determines a life table, once a family has been selected. Although technically a one parameter system, it could be argued that the choice of a family constitutes a separate dimension. The system was updated in 1989, primarily to include extensions of the model life tables to age 100+ [8].

The Ledermann system of model life tables was first published in 1959 and was subsequently revised over the course of the following decade [6]. This system is based on a factor analysis of some 157 empirical tables. The method of selection was less rigid than that of the Coale-Demeny tables, but they represent more developing country experiences. Analysis of the tables disclosed five factors that apparently explained a large proportion of the variability in mortality among the life tables. The extracted factors were related to 1.) the general level of mortality, 2.) the relation between child and adult mortality, 3.) mortality at older ages, 4.) mortality under age five, and 5.) the sex difference in mortality between the ages of 5 and 70 years. Ledermann then developed a series of one- and two- parameter model life tables based on these results.

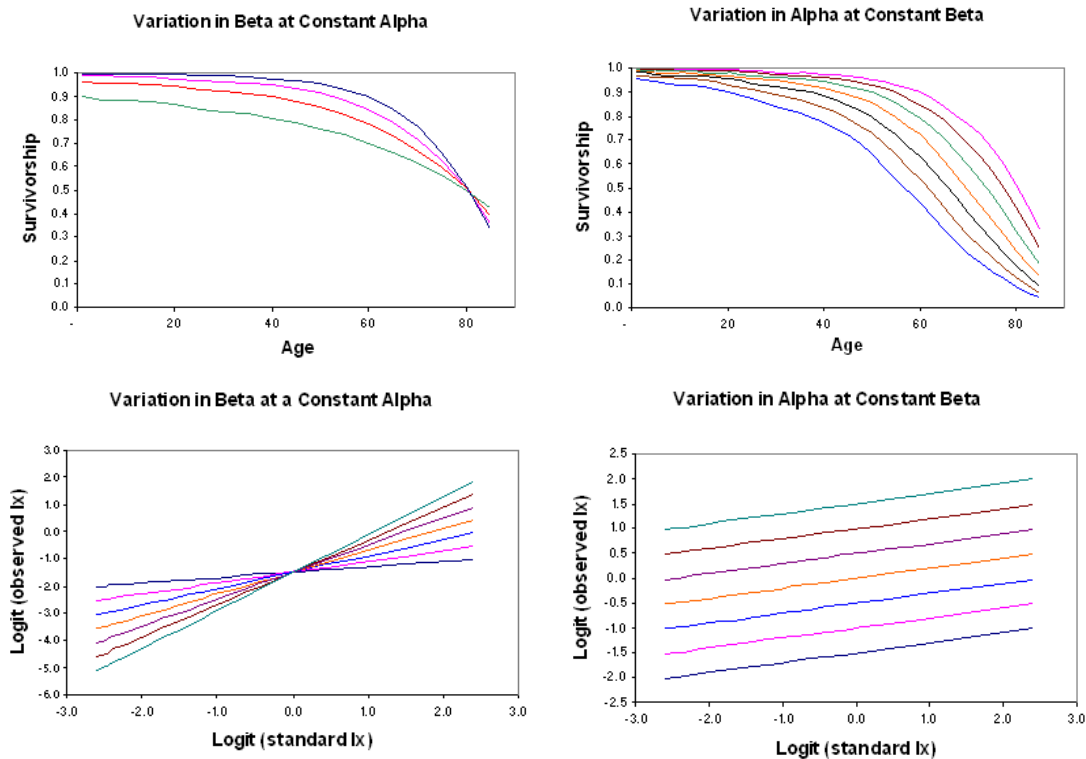
A different approach to constructing life table systems was first proposed by Brass in 1971 [2]. This relational system of life tables was built up from the observed structural relationship of survival curves among life tables. This system provides a greater degree of flexibility than the empirical models discussed above. It rests on the assumption that two distinct age-patterns of mortality can be related to each other by a linear relationship between the logit of their respective survivorship probabilities. Thus for any two observed series of survivorship values,  $l_x$  and  $l_x^s$ , where the latter is the standard, it is possible to find constants  $\alpha$  and  $\beta$  such that

$$\begin{aligned} \text{logit}(l_x) &= \alpha + \beta \text{logit}(l_x^s) \\ \text{if } \text{logit}(l_x) &= 0.5 \ln\left(\frac{(1.0 - l_x)}{l_x}\right) \\ \text{Then} \\ 0.5 \ln\left(\frac{(1.0 - l_x)}{l_x}\right) &= \alpha + 0.5\beta \ln\left(\frac{(1.0 - l_x^s)}{l_x^s}\right) \end{aligned}$$

for all age  $x$  between 1 and  $\omega$ . If the above equation holds for every pair of life tables, then any life table can be generated from a single standard life table by changing the pairs of  $(\alpha, \beta)$  values used. In reality, the assumption of linearity is only approximately satisfied by pairs of actual life tables. However, the approximation is close enough to warrant the use of the model to study and fit observed mortality schedules. The parameter  $\alpha$  varies the mortality *level* of the standard, while  $\beta$  varies the *slope* of the standard, i.e., it governs the relationship between the mortality in children and adults. Figure 1 shows the result of varying  $\alpha$  and  $\beta$ . As  $\beta$  decreases, there is higher survival in the older ages relative to the

standard, and vice versa. Higher values of  $\alpha$  at a fixed  $\beta$  lead to lower survival relative to the standard.

**Figure 1. Effect of Changing  $\alpha$  and  $\beta$  on Pattern of Observed Mortality Relative to Standard**



## Shortcomings of the Empirical Model Life Table Systems

There are three major criticisms of the original one-parameter UN model life tables. First, the fact that they are one-parameter systems makes them relatively inflexible. Such a single parameter model cannot adequately describe the complex mortality patterns available. In some cases, they have failed to describe adequately life tables that were known to be accurate [9]. Second, because the estimate of mortality in each age group is ultimately linked to the infant mortality rate through the chaining process, measurement errors are easily accentuated. The third criticism concerns the poverty of developing country life tables in the original design of the model. Additionally, some of the empirical tables included were of dubious quality.

The two-parameter UN model life tables for developing countries, while clearly an improvement over the one-parameter system, also suffer from some of these limitations. Perhaps the main criticism of the two-parameter UN system is that the strict selection criteria reduced the underlying set of empirical life tables to 72 (out of 286 originally retained). This relatively small number limits the applicability of the models to other populations. Moreover, the life tables are now outdated.

The Coale-Demeny model life tables had much higher standards of accuracy for the empirical tables. This demand, however, limited the number of non-European countries represented. As such, the Coale-Demeny tables may not cover patterns of mortality existing in the contemporary developing world. In fact, there are examples of well-documented mortality patterns that lie outside the range of the Coale-Demeny tables. In particular, Demeny and Shorter found no table within the family that adequately reflected the Turkish mortality experience [10]. The fact that one of the parameters of the Coale-Demeny system (the “family”) is discrete restricts the flexibility of the system, certainly in comparison to other systems where both parameters are continuous.

The Ledermann system is criticized primarily for its relative complexity which essentially precludes its use in most developing countries. Even though it does provide some flexibility through a wider variety of entry values, in practice most of these values are not easily estimated for most developing countries. This drawback reduces its relative advantages over the UN and the Coale-Demeny models. A second major limitation is that the independent variables used in deriving the model refer, with only one exception, to parameters obtained from data on both sexes combined. The user is, therefore, forced to accept the relationships between male and female mortality embodied in the model even when there is evidence to the contrary. For instance, it is near impossible to estimate a Ledermann model life table in which the male expectation of life exceeds that of females.

Another shortcoming common to all three empirical models is their dependence on the type of data that generated them. The datasets upon which they were built exclude a significant proportion of possible mortality schedules. Although the UN set of model life tables attempted to address this issue, there were serious flaws in the selection of life tables as well as the criteria of acceptance.

It is clear, therefore, that there are serious technical issues that complicate the use of existing empirical models in describing mortality patterns in contemporary developing countries. We are proposing a new modified two-parameter system of model life tables anchored on the logit system. The choice of the logit system was based on a careful comparative evaluation of the logit and the Coale-Demeny systems. This evaluation process is presented in a subsequent section.

## Modified Logit Life Table System

### *Principle of the Brass Logit Life Table System*

The Brass Logit life table system [2] belongs to a category of mortality models called relational models. It features a standard life table and two parameters which, through a mathematical transformation, relate the standard life table to any life table. The general shape of the survivorship functions is captured through the mortality standard while the parameters help to capture deviations from the standard.

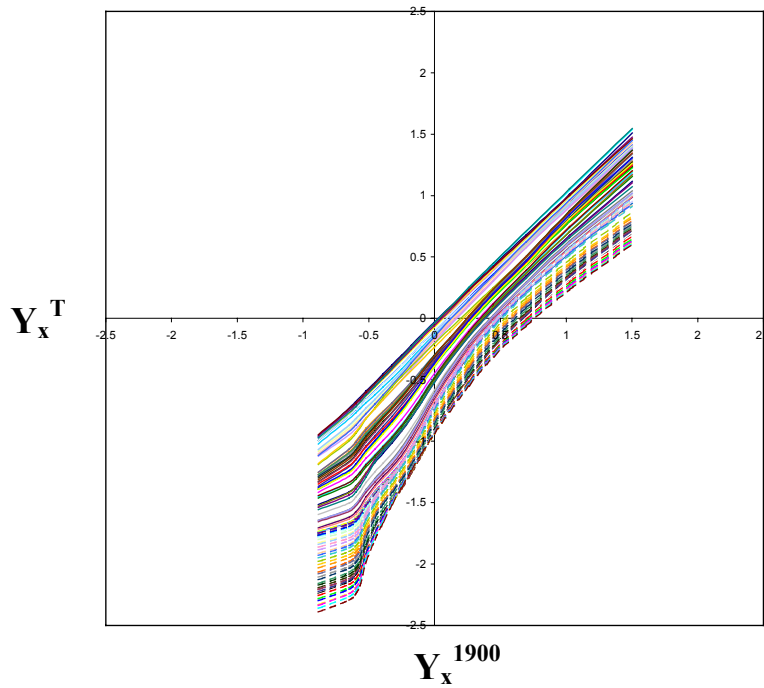
One problem with Brass' original relational model is that the relationship between two survivorship functions in logits is not always linear. Deviations from linearity appear to be particularly large when the observed mortality of a population is far from that of the standard. Thus the complexity of variations in levels and age patterns of mortality is not fully captured by the logit model. This observation led several authors to modify Brass' original model by including additional parameters that allow for bends in the survivorship function [11] [12]. This modification, however, is of little practical use, because the additional parameters are difficult to estimate empirically and complicate the applied use of the models.

Our modification of Brass' transformation is based upon some simple but powerful observations. The basic observation is that deviations from linearity follow some specific regularities which can be modelled in relation to the amount of mortality change between the standard and the observed life table.

These shifts in the structure of mortality can be illustrated by plotting a series of logit life table values against logit values taken from an earlier life table, and examining how the resulting curves depart from linearity. This is shown in Figure 2, which presents data for US males. In this figure, annual logit life table values from 1990 to 1995 are plotted against logit values for 1900, taken as a standard. It is clear that mortality change over time leads to a change in the age pattern of mortality that is not fully captured by the logit relational model. Indeed, if the logit transformation were fully appropriate, the successive plots in the figure would remain linear over time.



**Figure 2: Annual Logit Life Table Values (1900-95) vs. 1900 Logit Values (US Males)**



Our modification of Brass' logit transformation is based upon the observation that differences between observed and predicted logit values follow a pattern that is predictable as the mortality level of the observed life table deviates from that of the standard. That is, deviations from linearity in the original Brass model are linked to the relative difference between the mortality level of the standard and the mortality of the actual life table being estimated.

We can generalize the principle underlying Brass' approach to postulate that there is some transformation of the survivorship function such that all transformed survivorship functions are linear functions of each other. If this transformation can be identified then all survivorship functions can be expressed as a two-parameter transformation of all other survivorship functions. Formally:

$$\text{Eq. 1 } \Gamma(l_x) = \alpha + \beta \Gamma(l_x^s)$$

If the transformation  $\Gamma$  can be identified then all survivorship functions can be simply derived from the parameters  $\alpha$  and  $\beta$ . Brass' original proposal is that this transformation is a variant of a logit transformation such that:

$$\text{Eq. 2 } \Gamma(l_x) = 0.5 \ln\left(\frac{1-l_x}{l_x}\right)$$

The problem is that the logit transformation does not make most survivorship functions completely linear. To develop the modified model life table system, we have sought to identify a transformation that will do a better job of linearizing survivorship functions.

## Modifying the Logit Transformation

Several authors have modified the Brass Logit life table system by including extra parameters that allow for bends in the survivorship function [11] [12]. With more parameters the model life table system is more flexible but difficult in practice to use. Using the extensive dataset of high quality life tables for many countries at WHO, we have sought to try and identify alternative transformations,  $\Gamma$ , that would come closer to linearizing the relationship between most survivorship functions and thus having a two parameter model life table system with better predictive power. This search has been driven by empirical investigation of the differences between observed and predicted survivorship at each age. In this section, we provide the mathematical form of the transformation that we have found to give the best predictive power. The empirical fitting of the model is described below.

Comparison of observed and predicted age-specific mortality rates using the Brass logit transformation with a global standard reveals that as a population's mortality moves farther away from the standard systematic errors appear. Exploratory analysis revealed that this systematic error at each age was related to both the level of child mortality relative to the standard and the level of middle-age adult mortality relative to the standard. Based on this finding, a variety of alternative transformations have been investigated. Ultimately, based on multiple tests, the transformation that we have selected is:

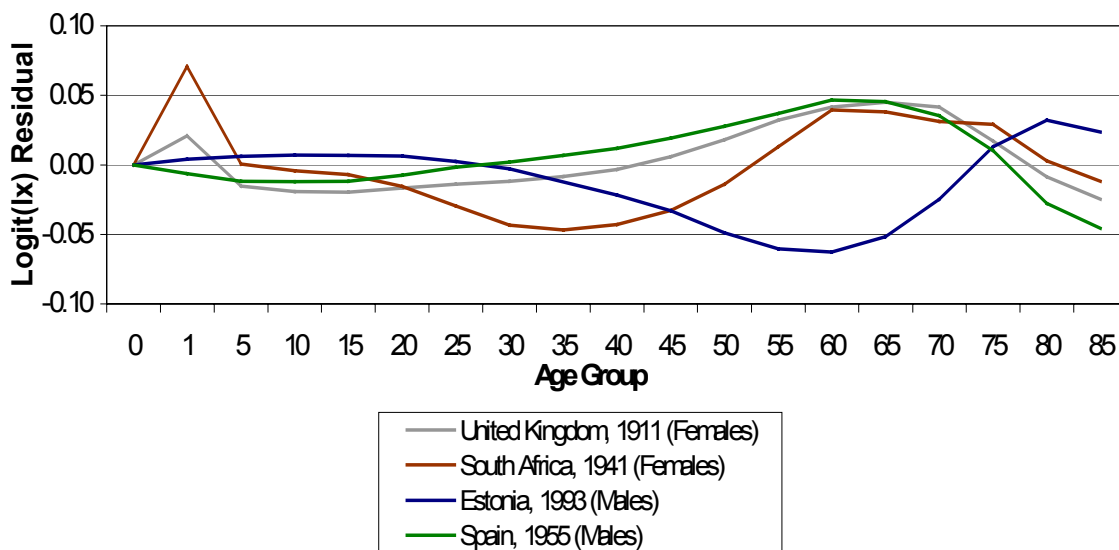
$$\text{Eq.3 } \Gamma(l_x) = \text{Logit}(l_x) - \gamma_x \left(1 - \left(\frac{\text{Logit}(l_5)}{\text{Logit}(l_5^s)}\right)\right) - \theta_x \left(1 - \left(\frac{\text{Logit}(l_{60})}{\text{Logit}(l_{60}^s)}\right)\right)$$

The parameter  $\gamma_x$  is a constant across populations with a different value for each age group as is the parameter  $\theta_x$ . The transformation is defined with respect to a single global standard. This standard is defined as the  $l_x$  values from the simple, unweighted average of all age-specific death rates from the 1802 populations retained in the final dataset. The set of  $\gamma_x$  and  $\theta_x$  are, accordingly, defined with reference to this global standard and are presented along with the parameter values in Table 3.

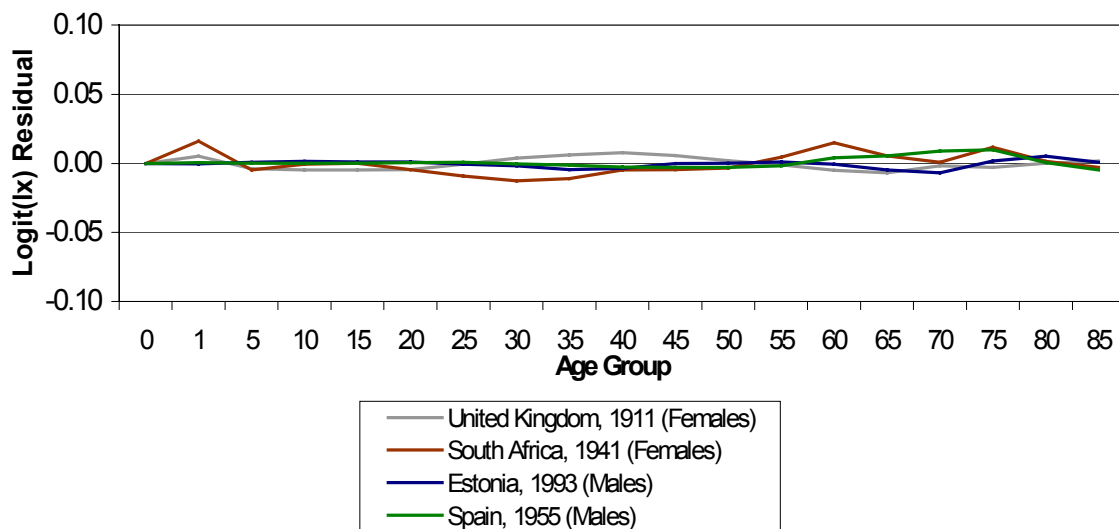
The effect of this transformation on a set of survivorship functions is shown in Figure 3. In the top panel, the deviations (residuals) by age between the logits of the observed  $l_x$  and those predicted from the original Brass system are plotted for four populations covering a range of mortality experiences. Substantial deviations are evident in all four populations, particularly at ages 0-4 and among older adults. In the bottom panel, the deviations based on this new transformation are shown for the same four populations. Clearly the fit is much better. Because this transformation makes the relationship between survivorship functions more linear with respect to age, a two parameter fit on the transformed standard will perform much better than the original simple logit transformation.

It is important to note that  $\gamma_x$  and  $\theta_x$  do not vary across countries or years. Because of this, each life table can still be uniquely defined with this transformation as a linear function of a standard using only two parameters. It is more advantageous to use the parameters of  $l_5$  and  $l_{60}$  to define a unique life table rather than  $\alpha$  and  $\beta$  since values of these are more readily interpretable than values of  $\alpha$  and  $\beta$ . It remains an empirical task to estimate the set of  $\gamma_x$  and  $\theta_x$  which is developed below.

**Figure 3a: Deviations Between Observed and Predicted Logits by Age Using the Original Brass Logit Model, Selected Countries.**



**Figure 3b: Deviations Between Observed and Predicted Logits by Age Using the Modified Logit Model, Selected Countries.**



## Dataset of Life Tables

The transformation selected to modify the original Brass Logit system includes three standard functions,  $l_x$ ,  $\theta_x$  and  $\gamma_x$  which are age- and sex-specific, but invariant across populations. These functions need to be estimated from a dataset of life tables which are considered to adequately reflect the age-sex patterns of mortality in as many countries as possible. Beginning in the 1960's, WHO began to systematically collect vital registration data on causes of death in countries, making every effort to complete the series back to 1950. For most countries, the most recent data refer to the period 1998-2000 [4]. The data for most countries contain the number of deaths by age, sex and cause, classified according to the Revision of the International Classification of Diseases in use. Data are collected by the conventional 5 - year age groups (0,1-4, 5-9, ..., 85+), although in recent years the terminal age group has been extended to 100+. For each year, mid-year population estimates by age and sex are also provided by reporting countries. These data have been screened for completeness using standard demographic tests, and only those country-years for which mortality was considered complete have been retained for this analysis.

This dataset was supplemented by life tables from two other sources. The historical compilation of life tables by Preston, Keyfitz and Schoen [13] were added to the data set for years not covered by the WHO mortality dataset. The mortality data underlying these life tables had been adjusted, where necessary, for under-reporting. To improve the coverage of developing countries in the dataset, the adjusted national life tables used by the United Nations [3] to produce their model life tables were also added.

Apart from the criteria of completeness and age- and sex-specific detail, we also applied criteria to exclude life tables of populations during periods of war or those affected by the Spanish influenza pandemic of 1918-19. Data for years prior to 1900 were excluded since the age patterns of mortality tended to be atypical. Small populations with a total size of less than one million people (both sexes combined) were also excluded to minimize the effects of random fluctuations in death rates.

The resulting set of 1,802 life tables used to develop the model are shown in Table 1. There is, of course, a preponderance of countries from Europe, North America and Australasia, but among the 63 countries represented, about one-third belong to developing regions. For several developed countries, historical datasets back to the beginning of the century have been included.

Table 2 summarises the characteristics of the life tables included in the dataset. The mean life expectancies are relatively high (67.5 years for males, 73.4 for females), reflecting the developed country bias, although the range of life expectancies (27 to 77 years for males, 29 to 84 years for females) certainly encompass the experience of all countries [4]. Average levels of child and adult mortality are not too dissimilar to what is observed in many developing countries today and again the range of values more than encompasses estimated levels across all developing countries, with the exception of a few countries in Africa (Namibia, Botswana, Zambia) where female mortality from HIV is extreme.

**Table 1: Life Tables Comprising the Empirical Dataset**

Country	Year(s)
Argentina	1966-70, 77-79, 82-97
Australia	1911, 1921, 1950-97
Austria	1955-99
Belarus	1981-98
Belgium	1954-98
Bangladesh (Matlab Region)	1975
Bulgaria	1964-98
Canada	1921, 1950-97
Chile	1909, 1920, 1930, 1940, 1950, 1955-82, 1984-98
Colombia	1960, 1964
Costa Rica	1956-83, 1985-98
Croatia	1982-98
Cuba	1970-98
Czech Republic	1934, 1982-99
Denmark	1921, 1930, 1952-98
El Salvador	1950, 1971
Estonia	1981-98
Finland	1952-98
France	1900-13, 1920-39, 1946-97
Georgia	1981-96
Germany	1969-98
Greece	1928, 1956-98
Guatemala	1961, 1964
Honduras	1961, 1974
Hungary	1955-99
India	1971
Iran	1974
Ireland	1950-98
Israel	1975-98
Italy	1901, 1910, 1921, 1931, 1951-97
Japan	1950-98
Korea, Rep. of	1973
Latvia	1980-98
Lithuania	1981-98

**Table 1: Life Tables Comprising the Empirical Dataset (continued)**

Country	Year(s)
Macedonia	1982-97
Mauritius	1990-98
Mexico	1958-59, 1969-73, 1981-83, 1985-98
Netherlands	1950-98
New Zealand	1901, 1911, 1950-98
Norway	1910, 1920, 1951-98
Panama	1960
Peru	1970
Philippines	1964, 1970
Poland	1959-98
Portugal	1920, 1930, 1940, 1955-98
Republic of Moldova	1981-98
Romania	1963, 1969-98
Russian Federation	1980-98
Singapore	1955-98
Slovakia	1982-98
Slovenia	1982-98
South Africa	1941, 1951, 1960
Spain	1930, 1940, 1951-69, 1971-98
Sri Lanka	1946, 1953
Sweden	1900-17, 1920-98
Switzerland	1951-98
Thailand	1970
Trinidad and Tobago	1990-97
Tunisia	1968
Ukraine	1981-98
United Kingdom	1901, 1911, 1921, 1931, 1950-98
United States of America	1900-16, 1920-41, 1945-98
Yugoslavia	1982-97

**Table 2: Characteristics of Life Tables Comprising the Empirical Dataset**

Sex	Parameter	Mean	Std. Dev.	Minimum	Maximum
Males:	$e_0$	67.46	6.16	26.64	77.29
	$5q_0$	0.039	0.047	0.005	0.439
	$45q_{15}$	0.208	0.076	0.087	0.762
	$20q_{60}$	0.636	0.078	0.422	0.906
Females:	$e_0$	73.39	6.81	29.20	84.00
	$5q_0$	0.033	0.043	0.003	0.427
	$45q_{15}$	0.121	0.066	0.049	0.656
	$20q_{60}$	0.478	0.099	0.222	0.833

## Empirical Fitting of the Modified Logit Life Table System

### A. Estimating $\theta_x$ and $\gamma_x$

By rewriting equations 3 and 1, we can express the age-specific  $\theta_x$  and  $\gamma_x$  and  $\alpha$  and  $\beta$  for each country (i)-year(j) in a way that we can estimate the parameter values using OLS regression:

$$\text{Eq.4 } \text{Logit}(l_{xij}) = \alpha_{ij} + \beta_{ij} \text{Logit}(l_x^s) - \gamma_x \left(1 - \left(\frac{\text{Logit}(l_{5ij})}{\text{Logit}(l_5^s)}\right)\right) - \theta_x \left(1 - \left(\frac{\text{Logit}(l_{60ij})}{\text{Logit}(l_{60}^s)}\right)\right)$$

The last two terms of Equation 4 are designed to control for the mortality differential between the standard life table and the actual life table. The first one captures differences in child mortality while the second captures differences in adult mortality up to age 60. We have estimated  $\alpha_{ij}$  and  $\beta_{ij}$  for each country-year life table and the set of  $\theta_x$  and  $\gamma_x$ , separately by sex, using OLS regression. The standard life table used is a sex-specific global standard calculated by taking the average of all sex-specific life tables included in the dataset. Due to computational limitations which prevented estimation of these parameters in one single regression using the entire dataset, we took 1000 random samples each of 100 life tables from the dataset and estimated  $\theta_x$  and  $\gamma_x$  in each case. As the typical deviation from the standard neither in the same direction nor of the same magnitude across age groups,  $\theta_x$  and  $\gamma_x$  are constants with different values for each age group. The estimates of  $\theta_x$  and  $\gamma_x$  were significant at the 95% confidence level for each age group except one, which was significant at the 90% level. The distributions of the resulting  $\theta_x$  and  $\gamma_x$  are summarised in Table 3 and Figure 4. Table 3 provides a summary of the estimated values at different percentiles of the distribution. These results are illustrated more fully in Figure 4 which presents standard box-whisker plots of the distribution of the estimates at each age. As the Figure shows, the age pattern for both parameters in males and females follows a consistent pattern. In all cases, we used the

50<sup>th</sup> percentile value of the distribution for  $\theta_x$  and  $\gamma_x$ , except for males at ages 65 and over for whom the 25<sup>th</sup> percentile values were used in order to improve the fit of the model at older ages.

### B. Estimating the models

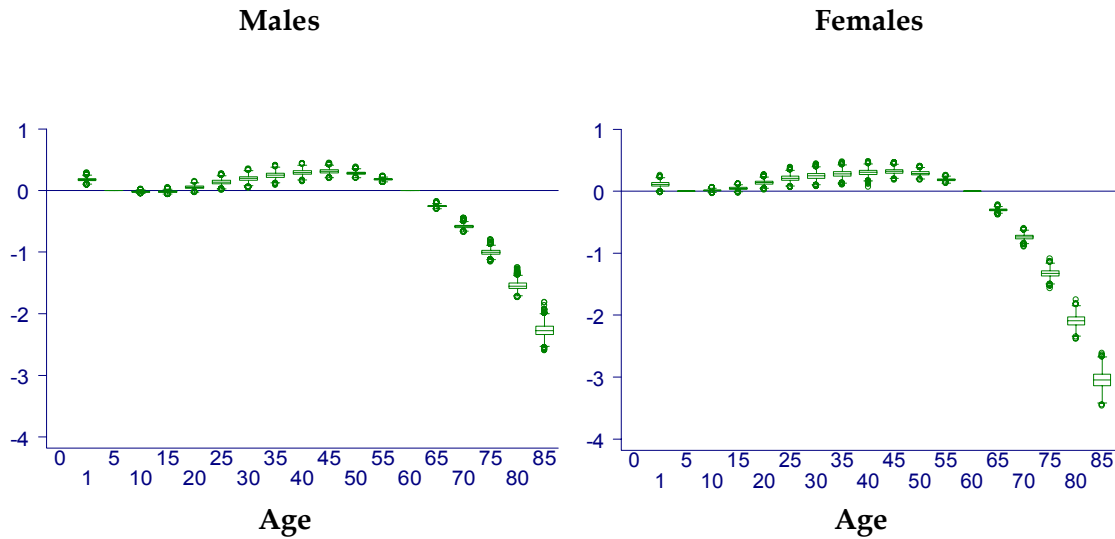
Having estimated  $\theta_x$  and  $\gamma_x$ , we can proceed to developing model life tables using the modified transformation. Firstly, we have constructed the global standard set of  $l_x$  values for males and females by simply amalgamating all countries in the dataset using the un-weighted average. Because the Modified Logit transformation includes two indices, one based on  $l_5$  and the other on  $l_{60}$ , which change as a survivorship function moves away from the global standard, it is more convenient to use  $l_5$  and  $l_{60}$  as the two parameters for the life table system. Any pair of  $l_5$  and  $l_{60}$  uniquely defines a life table because  $\alpha_{ij}$  and  $\beta_{ij}$  are a function of  $l_5$  and  $l_{60}$ . It can be shown that:

$$\text{Eq.5} \quad \alpha_{ij} = \frac{\text{logit} [l_5^j] \cdot \text{logit} [l_{60}^s] - \text{logit} [l_5^s] \cdot \text{logit} [l_{60}^j]}{\text{logit} [l_{60}^s] - \text{logit} [l_5^s]}$$

$$\text{Eq.6} \quad \beta_{ij} = \frac{\text{logit} [l_{60}^j] - \text{logit} [l_5^j]}{\text{logit} [l_{60}^s] - \text{logit} [l_5^s]}$$

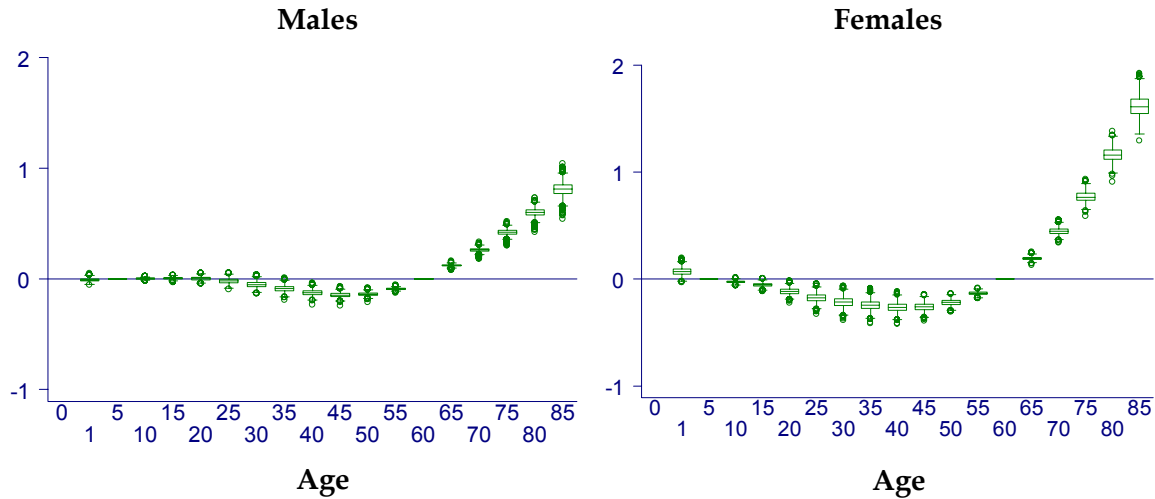
We proceed to use the transformation and these equations to generate model life tables for a wide range of combinations of  $l_5$  and  $l_{60}$ . We create the set of model life tables by sampling systematically from  $l_5$  and  $l_{60}$  and discarding combinations that are logically impossible (e.g.,  $l_5 < l_{60}$ ). Through this procedure, we generated approximately 50,000 life tables, each defined by a unique combination of  $l_5$  and  $l_{60}$ . The corresponding set of  $l_5$  and  $l_{60}$  values from these 50,000 life tables are shown in Figure 5. Since similar values of  $l_5$  and  $l_{60}$  imply implausibly low levels of  ${}_{45}q_{15}$ , there are no sample points close to the 45° axis.

**Figure 4a: Box-Whisker Plots of Estimated  $\gamma_x$  by Age**





**Figure 4b: Box-Whisker Plots of Estimated  $\theta_x$  by Age**



Using this collection of 50,000 life tables, we can visualize various life table parameters such as  ${}_nq_x$  and  $e_x$  in the two dimensional space defined by  $l_5$  and  $l_{60}$ . Figure 6a shows life expectancy at birth isoclines corresponding to given values of  $l_5$  and  $l_{60}$ . Each point on the isocline corresponds to a constant level of life expectancy generated by different age-patterns of mortality. The same life expectancy is possible with low child mortality and high adult mortality or higher child and lower adult mortality. The isoclines demonstrate that the same life expectancy can occur with widely varying age-patterns. This is illustrated more clearly in Figure 7 for an isocline with male life expectancy of 65 years. The figure shows the logarithm of death rates for four points on that isocline. The very substantial variation among the four sets of death rates illustrates the heterogeneity of age-patterns of mortality that we are trying to better capture with the Modified Logit life table system.

**Figure 5: Values of  $l_5$  and  $l_{60}$  from the Modified Logit System, Males**

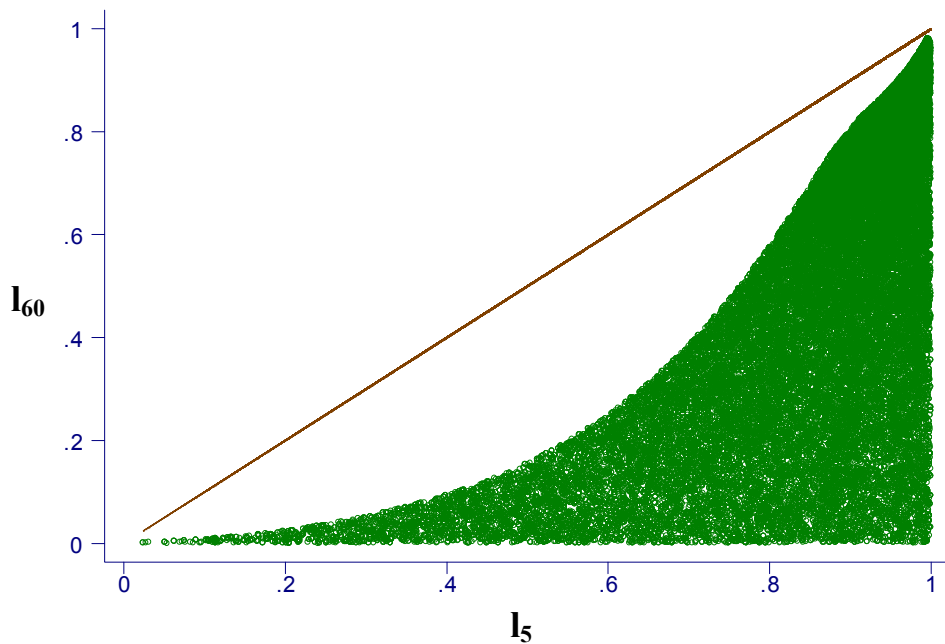


Figure 6: Isoclines of  $e_0$ ,  ${}_{45}q_{15}$  and  ${}_{20}q_{60}$ , Selected Values, Males

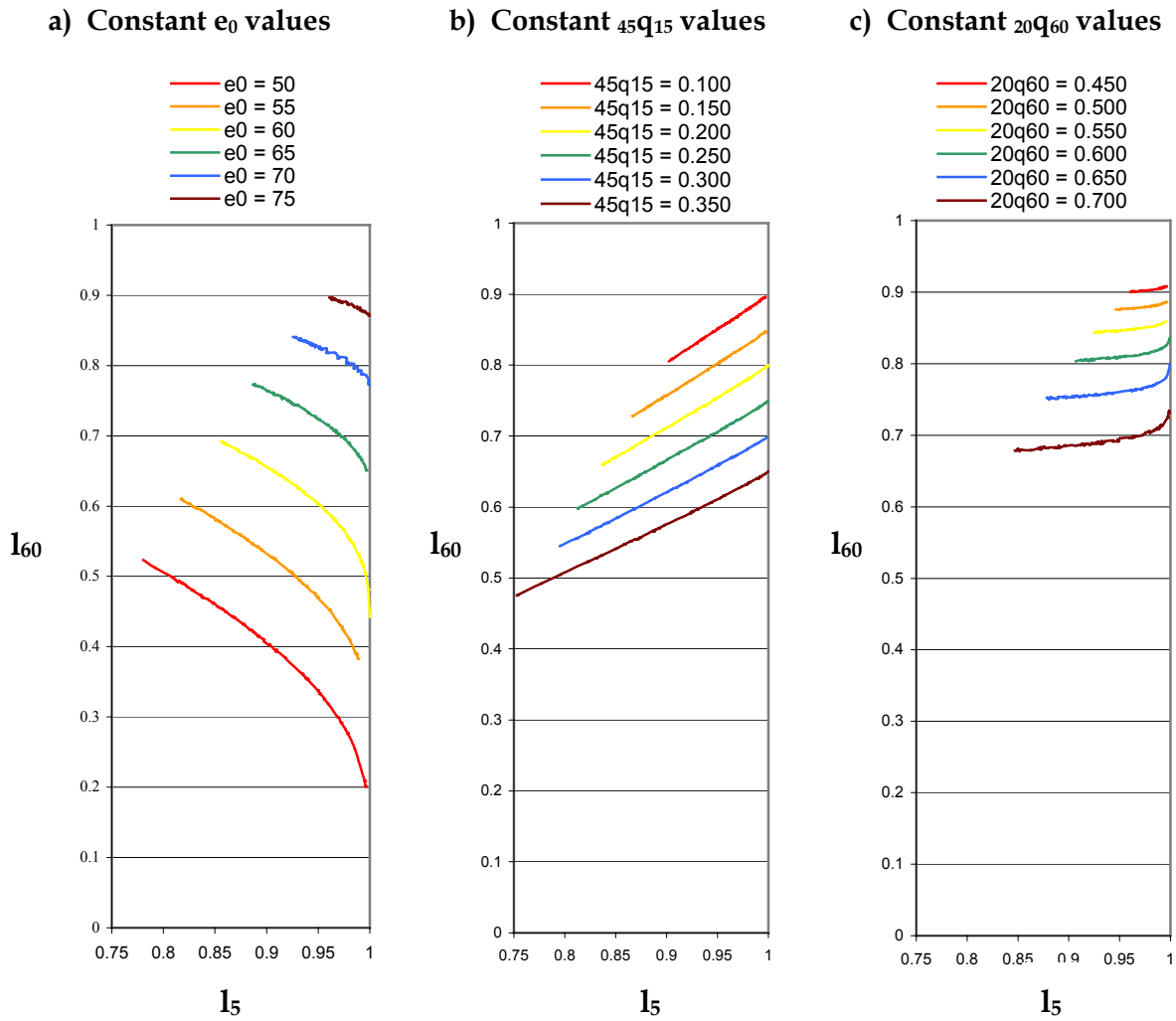
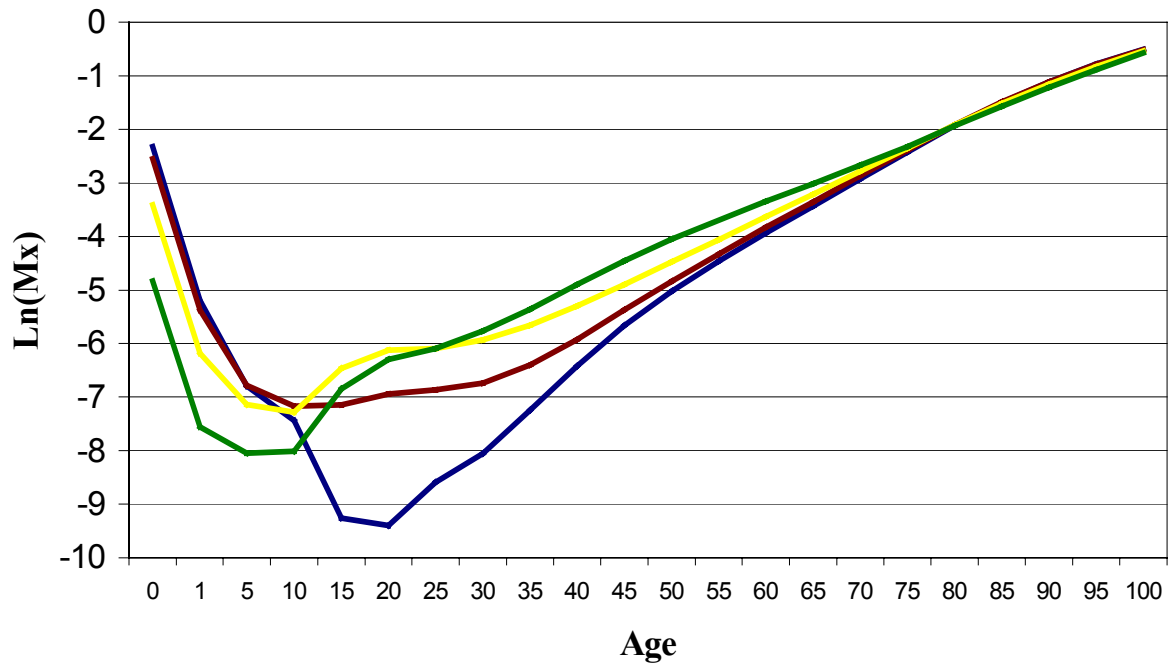


Figure 7: Log  $M_x$  for Four Populations with Male  $e_0 = 65$  Years



One can similarly visualize variation in other derivative life table measures based on the collection of 50,000 life tables according to levels of  $l_5$  and  $l_{60}$ . Figures 6b and 6c show, respectively, how adult mortality ( ${}_{45}q_{15}$ ) and mortality among the elderly ( ${}_{20}q_{60}$ ) vary according to the two index life table parameters,  $l_5$  and  $l_{60}$ . Consider Figure 6b. Each isocline corresponds to a fixed level of  ${}_{45}q_{15}$  which can be obtained from specified values of  $l_5$  and  $l_{60}$ . At a given level of child mortality, the gap between successive isoclines remains relatively constant, since mortality at ages 5-14 is low in all populations and relatively invariant. Figure 6c, on the other hand, indicates that the impact on older age mortality of declining levels of child and adult mortality is much less apparent, with only a modest overall level effect between child mortality and  ${}_{20}q_{60}$ . Rather, levels of  ${}_{20}q_{60}$  are much more strongly determined by levels of adult mortality, as one would expect.

A key advantage of the Modified Logit system is that any two life table parameters define a unique life table. The contour lines in Figures 6 demonstrate how there are an infinite set of combinations of  $l_5$  and  $l_{60}$  and thus complete life tables for any given level of a mortality measure such as life expectancy at birth. If two life table indices are known such as  ${}_{5}q_0$  and  $e_0$  then a unique life table is defined in this system at the point where the different contours intersect. For example, referring to Figure 6a, if we know  ${}_{5}q_0$  is 100 per 1000 and life expectancy at birth is 60 years, then the unique life table is defined by an  $l_5$  of 0.900 and an  $l_{60}$  of 0.652.

Because the contours cannot easily be defined analytically, we have developed two methods to identify the life table in the model life table system that matches any two life table parameters. Using the first method, we identify the full life table from a previously generated dataset of 50,000 life tables that provides the closest match to the estimated  ${}_{5}q_0$  and  ${}_{45}q_{15}$ . The second method requires generating a 10x10 matrix of life tables across the entire range of  $l_5/l_{60}$  space, whose parameters are then compared with the observed (or estimated) input values. After selecting the life table whose values of  ${}_{5}q_0$  and  ${}_{45}q_{15}$  match most closely with the inputs, a second 10x10 matrix of life tables is generated, centered on the values of  $l_5$  and  $l_{60}$  corresponding to the matching life table from the first set. The five-fold repetition of this process provides a degree of precision comparable to the first method and does not require the underlying dataset of 50,000 life tables.

**TABLE 3a: 2.5, 25, 50, 75 and 97.5 Percentile Values of  $\theta_x$  and  $\gamma_x$ , Males**

Standard	Age(x)	$\gamma_x$					$\theta_x$				
		Percentile					Percentile				
		2.5	25	50	75	97.5	2.5	25	50	75	97.5
1.0000	0	-	-	-	-	-	-	-	-	-	-
0.9695	1	0.1261	0.1583	0.1767	0.1948	0.2276	-0.0375	-0.0203	-0.0103	0.0002	0.0194
0.9612	5	0	0	0	0	0	0	0	0	0	0
0.9578	10	-0.0380	-0.0252	-0.0184	-0.0115	0.0029	-0.0101	-0.0024	0.0017	0.0056	0.0129
0.9550	15	-0.0469	-0.0272	-0.0171	-0.0067	0.0145	-0.0140	-0.0020	0.0038	0.0102	0.0209
0.9491	20	0.0009	0.0357	0.0538	0.0730	0.1079	-0.0284	-0.0094	0.0017	0.0128	0.0334
0.9404	25	0.0578	0.1084	0.1371	0.1631	0.2159	-0.0669	-0.0378	-0.0209	-0.0050	0.0252
0.9314	30	0.1049	0.1630	0.1966	0.2283	0.2890	-0.1028	-0.0712	-0.0519	-0.0338	0.0006
0.9209	35	0.1570	0.2161	0.2499	0.2818	0.3463	-0.1414	-0.1088	-0.0891	-0.0709	-0.0361
0.9072	40	0.2077	0.2618	0.2922	0.3227	0.3812	-0.1714	-0.1421	-0.1252	-0.1087	-0.0807
0.8879	45	0.2398	0.2861	0.3114	0.3365	0.3849	-0.1852	-0.1608	-0.1479	-0.1336	-0.1100
0.8598	50	0.2268	0.2625	0.2808	0.2997	0.3360	-0.1678	-0.1487	-0.1394	-0.1292	-0.1107
0.8185	55	0.1499	0.1704	0.1802	0.1907	0.2101	-0.1062	-0.0958	-0.0907	-0.0849	-0.0737
0.7592	60	0	0	0	0	0	0	0	0	0	0
0.6760	65	-0.2817	-0.2620	-0.2523	-0.2412	-0.2177	0.0991	0.1138	0.1195	0.1250	0.1350
0.5664	70	-0.6357	-0.6010	-0.5821	-0.5616	-0.5163	0.2174	0.2512	0.2625	0.2726	0.2925
0.4306	75	-1.0848	-1.0328	-1.0048	-0.9735	-0.9051	0.3601	0.4039	0.4210	0.4363	0.4672
0.2816	80	-1.6637	-1.5887	-1.5501	-1.5032	-1.3984	0.5131	0.5750	0.5991	0.6224	0.6664
0.1437	85	-2.4635	-2.3398	-2.2721	-2.2022	-2.0441	0.6729	0.7707	0.8094	0.8460	0.9185

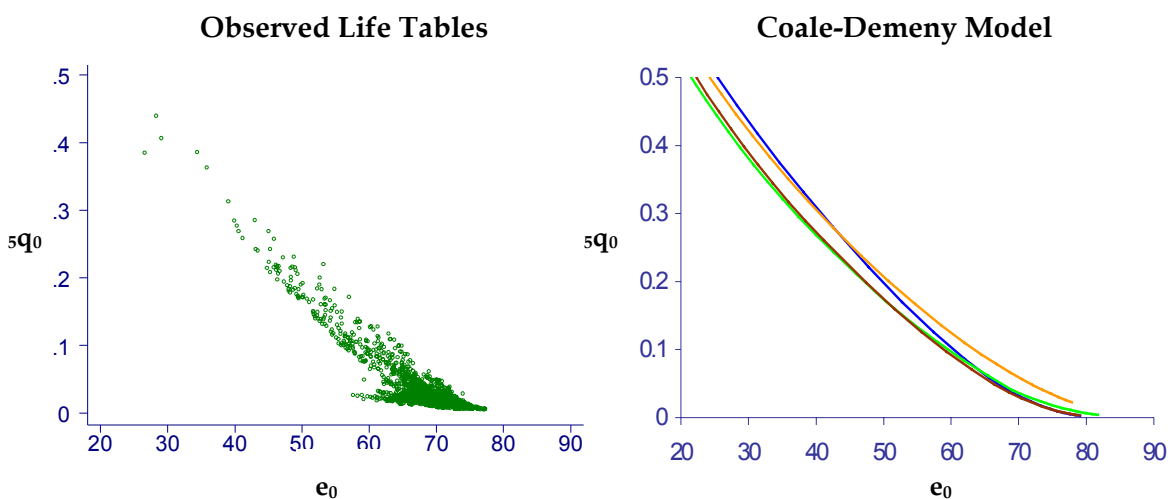
**TABLE 3b: 2.5, 25, 50, 75 and 97.5 Percentile Values of  $\theta_x$  and  $\gamma_x$ , Females**

Standard	Age(x)	$\gamma_x$					$\theta_x$				
		Percentile					Percentile				
		2.5	25	50	75	97.5	2.5	25	50	75	97.5
1.0000	<b>0</b>	-	-	-	-	-	-	-	-	-	-
0.9753	<b>1</b>	0.0216	0.0809	0.1077	0.1367	0.1919	0.0024	0.0459	0.0678	0.0919	0.1425
0.9675	<b>5</b>	0	0	0	0	0	0	0	0	0	0
0.9647	<b>10</b>	-0.0140	0.0039	0.0133	0.0224	0.0396	-0.0484	-0.0346	-0.0272	-0.0196	-0.0051
0.9626	<b>15</b>	0.0055	0.0321	0.0462	0.0615	0.0913	-0.0911	-0.0671	-0.0547	-0.0425	-0.0203
0.9591	<b>20</b>	0.0726	0.1136	0.1363	0.1598	0.2064	-0.1730	-0.1345	-0.1155	-0.0964	-0.0616
0.9547	<b>25</b>	0.1180	0.1763	0.2075	0.2408	0.3074	-0.2581	-0.2035	-0.1758	-0.1509	-0.1041
0.9496	<b>30</b>	0.1405	0.2116	0.2470	0.2839	0.3609	-0.3103	-0.2483	-0.2168	-0.1870	-0.1324
0.9434	<b>35</b>	0.1699	0.2423	0.2793	0.3158	0.3969	-0.3406	-0.2765	-0.2455	-0.2151	-0.1601
0.9352	<b>40</b>	0.2016	0.2694	0.3027	0.3384	0.4109	-0.3527	-0.2924	-0.2639	-0.2353	-0.1865
0.9239	<b>45</b>	0.2397	0.2905	0.3187	0.3480	0.4099	-0.3367	-0.2848	-0.2601	-0.2358	-0.1947
0.9075	<b>50</b>	0.2252	0.2654	0.2879	0.3117	0.3593	-0.2742	-0.2377	-0.2181	-0.2006	-0.1691
0.8835	<b>55</b>	0.1499	0.1732	0.1870	0.2000	0.2269	-0.1626	-0.1436	-0.1330	-0.1229	-0.1053
0.8488	<b>60</b>	0	0	0	0	0	0	0	0	0	0
0.7968	<b>65</b>	-0.3447	-0.3157	-0.3018	-0.2880	-0.2620	0.1613	0.1807	0.1913	0.2024	0.2231
0.7195	<b>70</b>	-0.8187	-0.7653	-0.7372	-0.7111	-0.6592	0.3894	0.4264	0.4462	0.4674	0.5066
0.6050	<b>75</b>	-1.4493	-1.3691	-1.3254	-1.2835	-1.2051	0.6803	0.7372	0.7658	0.8000	0.8587
0.4504	<b>80</b>	-2.2708	-2.1548	-2.0926	-2.0302	-1.9138	1.0335	1.1171	1.1586	1.2042	1.2865
0.2713	<b>85</b>	-3.3132	-3.1398	-3.0473	-2.9540	-2.7769	1.4257	1.5469	1.6083	1.6801	1.8066

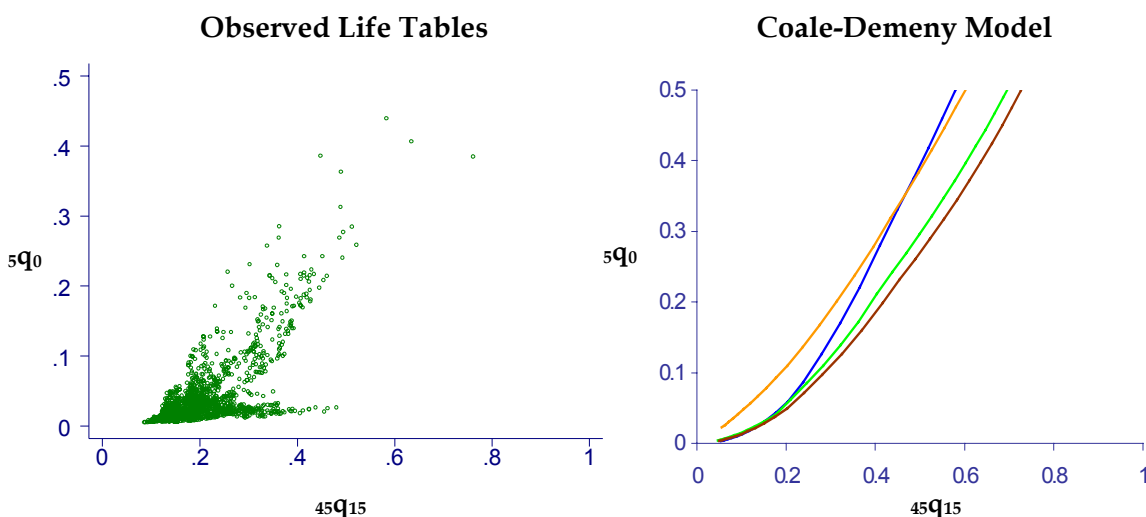
## Comparing the Coale-Demeny and Modified Logit Systems

How well do these model life table systems capture the observed range of mortality experience? As noted above, one important criterion for a model life table system is that it adequately represents the known range of mortality experience across countries. Figures 8-10 make three types of comparisons:  ${}_5q_0$  and  $e_0$ ,  ${}_5q_0$  and  ${}_{45}q_{15}$ , and  ${}_{45}q_{15}$  and  ${}_{20}q_{60}$ , respectively. In each figure the observed points from the underlying dataset are shown and compared with the Coale-Demeny model life table values. It is clear that the range of mortality experience captured in the Coale-Demeny system is much smaller than the observed range in the empirical life tables, particularly at medium levels of mortality.

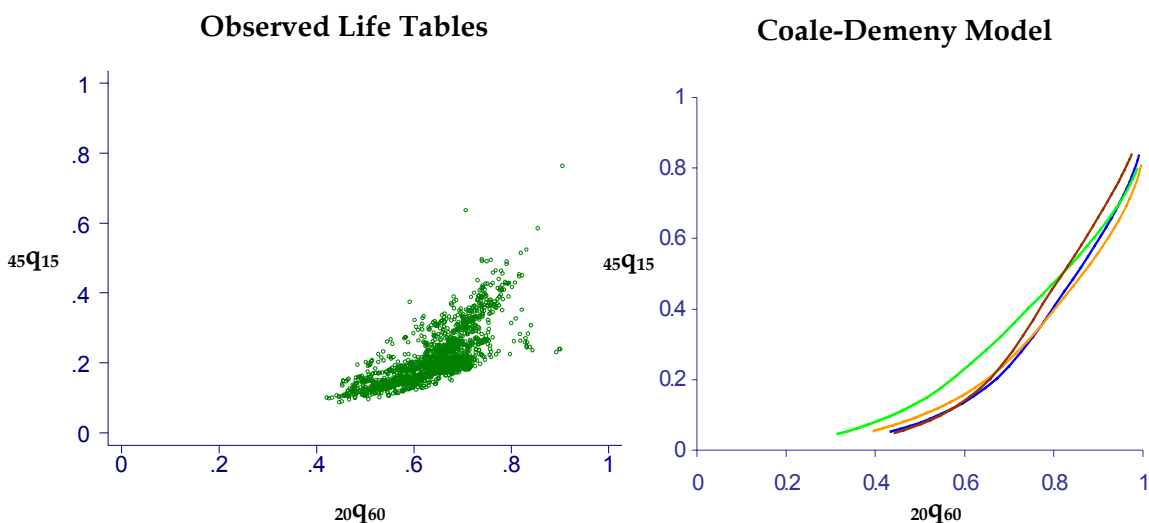
**Figure 8: Comparison of Observed Patterns of  ${}_5q_0$  and  $e_0$  vs. Coale-Demeny Model Values, Males**



**Figure 9: Comparison of Observed Patterns of  ${}_5q_0$  and  ${}_{45}q_{15}$  vs. Coale-Demeny Model Values, Males**



**Figure 10: Comparison of Observed Patterns of  ${}_{45}q_{15}$  and  ${}_{20}q_{60}$  vs. Coale-Demeny Model Values, Males**



The limited range of mortality patterns captured in the Coale-Demeny model life table systems can, in part, be explained by the emergence of the high adult mortality and low child mortality pattern now observed in parts of Eastern Europe and the Newly Independent States. When their system was developed little data was available on this pattern. Even excluding these countries, the range captured in this system is much smaller than the real variation seen worldwide.

In contrast, our Modified Logit life table system can capture the entire range of mortality patterns illustrated in Figures 8-10 as illustrated in the contour figures shown earlier. On this criterion, the Modified Logit system is clearly better able to capture the diverse array of mortality patterns now seen.

## Predictive Ability of the Modified Logit System

A key use of a model life table system is to create a full life table given information on only two life table indices such as life expectancy and child mortality or, more probably, adult mortality and child mortality. A strong test of this predictive use of a model life table system is to take empirical life tables, select a model life table using two aggregates from an empirical life table, and then compare the model life table age-specific death rates to the observed age-specific death rates. We have conducted two such tests for both model life table systems: choosing model life tables on the basis of  ${}_{5}q_0$  and  $e_0$ , and choosing on the basis of  ${}_{5}q_0$  and  ${}_{45}q_{15}$ .

First, for a random sample of 200 life tables from the life table dataset, we have used the Coale-Demeny and Modified Logit systems to select a model life table on the basis of  ${}_{5}q_0$  and  $e_0$ . The Coale-Demeny model has been selected by first matching each  $e_0$  on all families and then selecting the family with the closest  ${}_{5}q_0$ . The life table from the Modified Logit system has been selected using the matching algorithm described above. After repeating this procedure for each of the 200 selected life tables, the fit between

predicted and observed mortality rates has been summarized using the root mean square error in the log of the death rates, since the logarithm of the death rates allows a more meaningful comparison of death rates across age-groups.

Table 4 summarizes the goodness of fit statistics (root mean square error of the log death rates) from the two model life table systems. The upper panel gives the results for the first type of test described above where life tables were selected on the basis of  ${}_5q_0$  and  $e_0$ . As the Table clearly demonstrates, the Modified Logit system gives much better predictions of age-specific death rates than the Coale-Demeny system on the basis of this sample of 200 empirical life tables, particularly for females. Average root mean square errors from the Modified Logit system are typically about 60-65% of those from the Coale-Demeny system.

**Table 4: Comparison of Root Mean Square Error of  $\ln(m_x)$  from Fitting 200 Life Tables Using the Coale-Demeny and the Modified Logit Systems**

Life Tables Selected on the Basis of $e_0$ and ${}_5q_0$ :		Coale-Demeny	Modified Logit
	<i>Males</i>	0.3282	0.2156
	<i>Females</i>	0.2476	0.1464
Life Tables Selected on the Basis of ${}_{45}q_{15}$ and ${}_5q_0$ :			
	<i>Males</i>	0.4042	0.1975
	<i>Females</i>	0.1846	0.1207

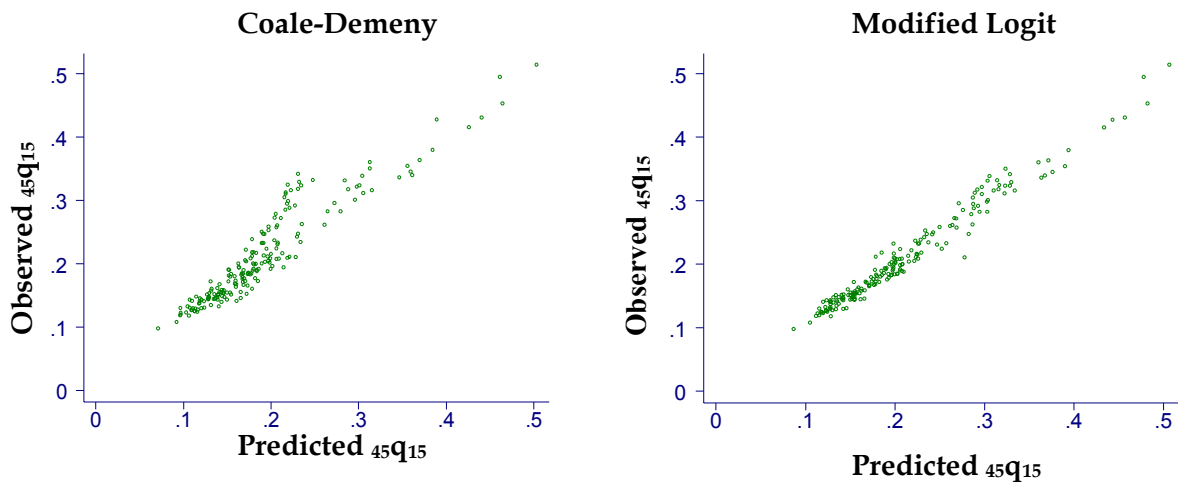
The second test that we have used to assess the predictive power of these systems is to select model life tables on the basis of  ${}_5q_0$  and  ${}_{45}q_{15}$ , a situation that is likely to be more commonly encountered. This is a more difficult test as the selection of the model life table is based on indices of mortality that cover a smaller age-range than life expectancy at birth. For each observed life table in the random sample of 200 life tables from the life table dataset, the Coale-Demeny model life table has been selected by matching on  ${}_{45}q_{15}$  and then choosing among the families on the basis of the  ${}_5q_0$ . The life table from the Modified Logit system has been selected by matching on the  ${}_5q_0$  and  ${}_{45}q_{15}$ . The predicted age-specific death rates have again been assessed using the root mean square error for the log death rates. Again, the Modified Logit system clearly outperforms the Coale-Demeny system, with average root mean square errors being about 50% of those from the Coale-Demeny system for males, and about one-third lower for females. This sex differential in relative performance of the two approaches may relate to the fact that the variance in adult male mortality is greater than for females. Hence, methods which reduce deviations in death rates at older ages will be of greater relevance to males where variability is greatest.

Figures 11 and 12 show the relative performance of the two model life table systems in predicting the actual observed probability of adult death ( ${}_{45}q_{15}$ ) (Figure 11) and life expectancy at birth (Figure 12) based on a random sample of 200 life tables from the empirical data set. These are two life table parameters which any model life table system would be most often required to predict. The observed values of  $l_5$  and  $l_{60}$  were input into the model and the full life table predicted. If a system could exactly predict the true

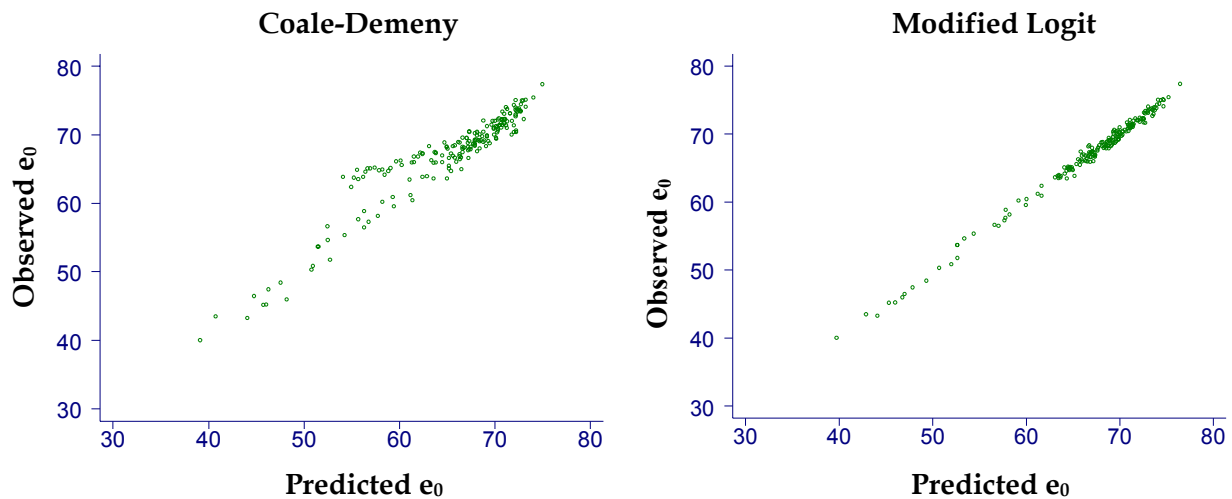


life table values, then all sample points would lie on a straight line. As Figure 11 illustrates, the Modified Logit system more successfully predicts the true probability of adult death (for males) than the closest match from the Coale-Demeny system, selected on the basis of  ${}_{5}q_0$  and  $e_0$ , as described earlier. In particular, the Coale-Demeny system performs relatively poorly for true levels of  ${}_{45}q_{15}$  in excess of about 150 per 1000, which would include much of the contemporary developing world. A similar pattern is apparent from Figure 12 which clearly shows the much closer fit between observed and predicted male  $e_0$  for this sample of countries compared with the Coale-Demeny system, selected on the basis of  ${}_{5}q_0$  and  ${}_{45}q_{15}$ , irrespective of the level of true life expectancy.

**Figure 11: Observed vs. Predicted Male  ${}_{45}q_{15}$  Using the Coale-Demeny and Modified Logit Systems, Selecting on the Basis of  ${}_{5}q_0$  and  $e_0$  (n=200)**



**Figure 12: Observed vs. Predicted Male  $e_0$  Using the Coale-Demeny and Modified Logit Systems, Selecting on the Basis of  ${}_{5}q_0$  and  ${}_{45}q_{15}$  (n=200)**



In addition to assessing the overall fit between predicted age-specific death rates and those actually observed, we have tested for any systematic bias in the death rates at different ages. Table 5a summarizes the regression results of the observed on predicted values for various life table parameters, using the 25<sup>th</sup> and 50<sup>th</sup> percentiles of the distribution of  $\theta_x$  and  $\gamma_x$  values. If the Modified Logit system were able to perfectly predict the observed life table parameter (e.g.,  ${}_{45}q_{15}$  or  ${}_{20}q_{60}$ ), then the coefficient of the regression would equal one and the constant would be zero. As is clear from Table 5a, this is very nearly the case for all tests conducted on the entire data set of life tables, with the greatest departures from unity at ages 60-80 years and at 15-60 years for females. For females, the 50<sup>th</sup> percentile performed about as well (ages 60 to 80) or substantially better (ages 15 to 60) in terms of the coefficient, and hence we have opted to use the median value of the distribution for women. For males, however, while the 50<sup>th</sup> percentile leads to a better fit on observed death rates at ages 15 to 60, it does much less well at ages 60 to 80. As a result, we have used the 50<sup>th</sup> percentile of the distribution at all ages below 65 and the 25<sup>th</sup> percentile values for  $\theta_x$  and  $\gamma_x$  at all ages 65 and over.

Table 5b summarizes the regression results where the predicted death rates have been estimated from this age-mix of  $\theta_x$  and  $\gamma_x$  for males. The value of the coefficient is closer to unity at ages 60 to 80 and the error in the constant has also been reduced. The reduction in the bias of the predicted values of probability of death at higher ages had little or no effect on the overall  $R^2$  for  $e_0$  which was very close to 1.

**Table 5a: Results of Regression of Selected Observed Life Table Parameters on Those Predicted by the Modified Logit System (n=1802)**

	Males (25 <sup>th</sup> Percentile $\gamma_x$ and $\theta_x$ )				Males (50 <sup>th</sup> Percentile $\gamma_x$ and $\theta_x$ )			
	$\alpha$	$\beta$	$R^2$	RMSE	$\alpha$	$\beta$	$R^2$	RMSE
$E_0$	-0.550238	1.008225	0.9995	0.1381	-0.432490	1.006685	0.9995	0.1383
${}_{45}q_{15}$	-0.006151	1.036309	0.9678	0.0136	0.004351	0.980468	0.9742	0.0122
${}_{20}q_{60}$	-0.032150	1.048287	0.9620	0.0151	-0.055391	1.085264	0.9686	0.0137

	Females (25 <sup>th</sup> Percentile $\gamma_x$ and $\theta_x$ )				Females (50 <sup>th</sup> Percentile $\gamma_x$ and $\theta_x$ )			
	$\alpha$	$\beta$	$R^2$	RMSE	$\alpha$	$\beta$	$R^2$	RMSE
$E_0$	-0.292568	1.003791	0.9995	0.1576	-0.334564	1.004834	0.9995	0.1580
${}_{45}q_{15}$	-0.009056	1.091761	0.9789	0.0097	0.001337	0.990415	0.9855	0.0080
${}_{20}q_{60}$	0.006349	0.985047	0.9802	0.0139	-0.016185	1.033520	0.9850	0.0121

**Table 5b: Results of Regression of Selected Observed Life Table Parameters on Those Predicted by the Modified Logit System Using Mixed  $\gamma_x$  and  $\theta_x$  Values**

	Males (Mixed Percentile $\gamma_x$ and $\theta_x$ )*			
	$\alpha$	$\beta$	R <sup>2</sup>	RMSE
$e_0$	0.979177	0.985246	0.9994	0.1525
${}_{45}q_{15}$	0.004351	0.980468	0.9742	0.0122
${}_{20}q_{60}$	0.017233	0.972280	0.9481	0.0177

\* Using 50<sup>th</sup> for ages<65 and 25<sup>th</sup> for ages>=65 (Males only)

## Discussion

In this paper, we have demonstrated that the Modified Logit life table system using a single global standard can represent the full range of mortality patterns seen across the high quality life tables available internationally. This two-parameter model life table system also generates better predictions of age-specific mortality rates than the Coale-Demeny system. Parameterizing this system using  $l_5$  and  $l_{60}$  also makes the parameters easier to understand. Another advantage of this system is that it requires a limited number of calculations, and it is thus easy to implement. Along with the three sets of parameters estimated here ( $l_x$ ,  $\gamma_x$  and  $\theta_x$ ), the only required data are two empirical life table values,  $l_5$  and  $l_{60}$ . The full life table can then be easily estimated.

The main limitation of this model life table system and the tests of its predictive power is that the sample of high quality life tables is heavily weighted towards populations with life expectancies between 60 and 73 (for males) and 66 and 80 (for females). The addition of more high quality and recent life tables for high mortality populations might require further modifications to the life table system. While these would be highly desirable, it is unlikely that they would alter the predictive ability of the system.

It is also uncertain how the model system would perform in countries with high levels of HIV. It is quite possible that in high HIV settings the age pattern of mortality projected out of sample by the model may not be accurate, although this cannot be tested due to the lack of high quality life tables for these countries. For the present, the model can and has been used to estimate life tables in the absence of HIV, with HIV death rates then added on *a posteriori* [4]. A comparison of the life tables for three high-HIV mortality countries (Zimbabwe, South Africa, and Tanzania) with the predictions from the model, using the  $l_5$  and  $l_{60}$  values from these life tables, suggests that the model can reliably reproduce life tables estimated in this fashion. In most cases, predicted life expectancy at birth was within 0.5 years of the value estimated from this two-stage procedure, with an even closer agreement for levels of adult mortality.

From our empirical fitting of the model to observed data, it is clear that there remains a slight bias in old age mortality where the modified system has not succeeded in fully

compensating for the lack of linearity mentioned earlier. Further work is required to attempt to improve the model to reduce these deviations at older ages.

The use of the Coale-Demeny and UN systems is so widespread in demographic estimation that there are often circular arguments about levels and patterns of adult mortality. One set of analysts often use the results of other demographic analyses founded on these model life table systems without realizing that they substantially underestimate the variation in age-specific mortality patterns seen in the real world. The use of models is so deeply embedded in available international datasets that it can be difficult to formulate real empirical tests of these models. We have tried to ensure that the observed life tables used in this analysis have not been modified using model life table systems, and hence that the modified system is based exclusively on observed data.

One implication of this analysis is that for Sub-Saharan Africa in particular there is much more uncertainty about levels of adult mortality than implied in currently available demographic estimates such as the UN Population Division life tables [14]. Often, levels of adult mortality have been estimated by selecting a life table on the basis of estimated child mortality and an arbitrary choice of a model life table family (often West by default). This has tended towards a one to one mapping of child mortality to adult mortality prior to the HIV epidemic. In reality, even the empirical record of countries outside Africa suggests that there can be much greater variation in levels of adult mortality as compared to child mortality than captured in the Coale-Demeny and UN model life tables.

The Modified Logit life table system which we have proposed appears to provide the best possible estimates for countries without good vital registration. Yet the model depends on the availability of reliable estimates of child and adult mortality. While these generally exist for child deaths [15], there are more than 60 countries, mostly in Sub-Saharan Africa, where levels of adult mortality are unknown or known poorly from indirect methods [4]. There is an urgent need to better measure levels of adult mortality directly in these countries in order to better guide public policy for health improvement.

## References

1. Coale A and Demeny P (1966). Regional Model Life Tables and Stable Populations. Princeton: Princeton University Press.
2. Brass, W (1971). On the scale of mortality. In W. Brass (ed.). Biological Aspects of Demography. London: Taylor and Francis.
3. United Nations (1981). Model Life Tables for Developing Countries. New York: United Nations. (Sales No. E.81.XIII.7)
4. Lopez AD, Ahmad OB, Guillot M, Inoue M, Ferguson BD (2001). Life Tables for 191 Countries for 2000: Data, Methods and Results. Geneva, World Health Organization (GPE Discussion Paper No. 40)
5. United Nations (1983). Manual X: Indirect Techniques for Demographic Estimation. New York: United Nations. (Sales No. E.83.XIII.2)
6. Ledermann, S. (1969). Nouvelles tables-type de mortalité: travaux et document. Paris: Institut National d'Etudes Demographiques.
7. United Nations (1955). Age and sex patterns of mortality: model life tables for under-developed countries. New York: United Nations. (Sales No. 1955.XIII.9)
8. Coale A and Guo G (1989). Revised regional model life tables at very low levels of mortality. *Population Index*. 55(4): 613-643.
9. Menken, J. (1977). Current status of demographic models. *Population Bulletin of the United Nations*. 9: 22-34.
10. Demeny P and Shorter F (1968). Estimating Turkish mortality, fertility and age structure. Statistics Institute, Istanbul University.
11. Zaba, B. (1979). The four-parameter logit life table system. *Population Studies*. 33: 79-100.
12. Ewbank D, Gomez de Leon J, Stoto M. (1983). A reducible four-parameter life table system. *Population Studies*. 37: 105-127.
13. Preston SH, Keyfitz N and Schoen R. (1972). Causes of Death: Life Tables for National Populations. New York: Academic Press.

14. United Nations (2001). The 2000 Demographic Assessment. New York: United Nations.
15. Ahmad OB, Lopez AD, Inoue M (2000). The decline in child mortality: a reappraisal. Bulletin of the World Health Organization. 78: 1175-1191.