International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

Age-Period-Cohort (APC) Analysis of Mortality with Applications to Soviet Data

Willekens, F. and Scherbov, S.

IIASA Working Paper

WP-91-042

1991

Working Paper

Age-Period-Cohort (APC) Analysis of Mortality with Applications to Soviet Data

F. Willekens and S. Scherbov

WP-91-42

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg Austria

Telephone: +43 2236 715210 □ Telex: 079137 iiasa a □ Telefax: +43 2236 71313

AGE-PERIOD-COHORT (APC) ANALYSIS OF MORTALITY WITH APPLICATIONS TO SOVIET DATA

F. Willekens and S. Scherbov

WP-91-42
November 1991

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg Austria

Telephone: +43 2236 715210 □ Telex: 079137 iiasa a □ Telefax: +43 2236 71313

ABOUT THE AUTHORS

Frans Willekens is the Director of the Population Research Centre, University of Groningen, P.O. Box 800, NL-9700 AV Groningen, The Netherlands.

Sergei Scherbov is a Research Scholar in the Population Program at the International Institute for Applied Systems Analysis, and is currently on the staff at the Population Research Centre, University of Groningen, The Netherlands.

ABSTRACT

The paper consists of two parts. The first part reviews the research on age-period-cohort analysis (APC) of mortality where APC models are extensively applied. A number of solutions to the identification problem are reviewed. It is claimed that the identification problem is not a problem of model specification, but a problem of measurement and specification.

In the second part of the paper, APC models are applied to mortality data. With the recent opening of the Soviet society, many demographic data that have been inaccessible for researchers have now become available. This is especially true of mortality data. By applying APC analysis for age-specific mortality rates for Soviet republics (or former Soviet republics), the authors try to separate contemporary and historical factors and thus capture several events that took place in Soviet history. Comparative analysis of age, period and cohort effects for different regions of the USSR are presented.

TABLE OF CONTENTS

AGE-PERIOD-COHORT (APC) ANALYSIS OF MORTALITY WITH APPLICATIONS TO SOVIET DATA

Frans Willekens and Sergei Scherbov

1. INTRODUCTION

Classical cohort analysis, or age-period-cohort (APC) analysis, is a method for exploring time series of demographic data and for the comparison of life courses of different cohorts. The time series generally consist of data classified by age and period (e.g. calendar year). Sometimes data are grouped by age and cohort. Variations in the age profiles are attributed to contemporary and historical factors. The contemporary factors are usually referred to as 'period effects' and are generally approximated by the calendar year. The historical factors represent the influence of the past on current behavior or experience and are usually referred to as 'cohort effects'. Cohort effects occur whenever the past history of individuals exerts an influence on their current experience or behavior in a way that is not fully captured by the age variable (Hobcraft et al., 1985, p. 97). The main contribution of APC analysis is that the impact of societal and technological processes on demographic experience is conceptualized in its historical and contemporary dimension.

The cohort or generation is an important concept in the study of changes in human behavior and experiences over time. The interest in cohort analysis is particularly large when discontinuities occur in trends. Cohort analysis is expected to reveal and quantify the impact in time of these discontinuities. Mannhein, who introduced the cohort concept into sociology in 1928, ascribed growing interest in the cohort problem to political discontinuities in the late 19th century. The trends that are studied may relate to social, economic, demographic, health or other variables. As a consequence, cohort analysis is broadly applied. For a general review covering several disciplines, see Hastings and Berry (1979). Hobcraft et al. (1982) review demographic studies, Breslow (1985) and Lidell (1985) discuss cohort analysis in epidemiology, Baltes et al. (1979) address cohort studies in psychology and Attias-Dunfot (1988) presents a comprehensive treatment of the cohort (generation) concept and generation theories in sociology (with at least one major omission, namely the classical article by Ryder published in 1965).

The subject of this paper is the method of age-period-cohort analysis and its application to mortality trends in the Soviet Union. Section 2 presents a general introduction to APC analysis. The literature on APC analysis of mortality data is reviewed in Section 3. The main part of the paper (Section 4) deals with the specification and estimation of the APC model and the interpretation of the model parameters, which represent the effects of age, period and cohort. In Section 5, the model is applied to unravel age, period and cohort effects in mortality data of the USSR and its regions. The data consist of age-specific mortality rates of the USSR from 1958 and of its regions from 1970. The results are compared to those of a recent study of trends in Soviet mortality. Section 6 concludes the paper.

2. APC ANALYSIS: AN INTRODUCTION

In traditional APC analysis, the contemporary factors are approximated by the current period and the historical factors are represented by the year or period of birth. Current period and period of birth are not causal factors in the analysis. They are crude indications of the macro-setting that changes over time and in which demographic phenomena are embedded. In the traditional analysis, the demographic rates, measured for a given age group during a given period, are decomposed into an effect of age grouping (age effect), an effect of contemporary factors (period effect) and lasting effect of historical factors experienced by the group of people to which the rate applies (cohort effect). A (birth) cohort is generally defined as a group of people born during the same period; in APC analysis, it is interpreted as a group of people who lived through comparable historical or structural contexts (e.g. depression, war period, period of rapid technological change). They may be referred to as 'contemporaries'. Although the impact of past common experiences remaining at the time of observation is likely to differ for each member of the group, there is probably some effect that is still felt by all members of the group. That effect is the cohort effect. APC analysis attempts to unravel inter-cohort differences and intra-cohort variations.

The APC analysis combines the two viewpoints traditionally distinguished by a demographer when analyzing demographic data. One approach examines changes from year to year. Period analysis, as this approach is known, is particularly useful when rapid changes occur, such as technological or legal changes that directly affect the controllability of demographic processes, or a war or a revolution resulting in transitory behavioral changes such as the postponement of births. The other approach, cohort analysis, is better suited for the study of fundamental changes in behavior such as an increase in health conditions and life expectancy. For a comprehensive treatment of APC analysis in demographic and social research, see Mason and Fienberg (1985).

The traditional APC model is not an explanatory model but a statistical accounting scheme. To interpret the period and cohort effects, one must look for attributes of the historical contexts that brought about the effects; the age effects must be related to attributes of human development over the life-span. The new approach to cohort or APC analysis introduces two major changes. First, it adopts a multilevel perspective: the characteristics of a cohort are aggregated outcomes of the individual behavior of cohort members in the societal and technological contexts. In other words, the effects of contemporary and historical factors on demographic change are <u>mediated</u> by individual characteristics, including the stage in the life course. Second, it adopts a <u>process perspective</u> and calls for longitudinal data to investigate the processes as they evolve.

> Nowadays, dying is rarely a sudden event; it is usually the culmination of a lengthy process during which the individual has suffered to a greater or lesser degree from diseases or handicaps which affect his mortality risk. It is thus a complex process (morbidity), the conclusion of which (death) cannot be studied

> without taking into account the process which preceded it: the population distribution of morbidity is a prime determinant of mortality risks and in turn the selection effects of mortality determine who survives with a chronic degenerative disease (van Poppel, 1990, p. 241).

The modern APC analysis is process-oriented and integrates life course analysis into cohort analysis. The integration signifies that events are simultaneously studied in two time scales: age and historical time. Representatives of this new school of thought include the sociologist Mayer (see e.g. Mayer and Huinink, 1990) and the demographers Caselli and Wunsch (see e.g. Caselli et al., 1990).

Caselli and Wunsch develop their theoretical framework as part of a study of adult mortality. It is interesting to note that the 'modern' approach adopts more fully than previous approaches the version of cohort theory promoted by Ryder (1965) in his classic paper. Ryder states that

> transformations of the social world modify people of different ages in different ways; the effects of these transformations are persistent. In this way a cohort meaning is implanted in the age-time specification. Two broad orientations for theory and research flow from this position: first, the study of intra-cohort development throughout the life-cycle; second, study of comparative cohort careers, i.e. intercohort temporal differentiation in the various parameters that may be used to characterize these aggregate histories (Ryder, 1965, p. 861).

Ryder emphasizes the need for a theoretical formulation of the phenomena under study and a focus on processes instead of on "the illusion of immutable structure" (Ryder, 1965, p. 859). The approach advocated by Ryder is similar to the one suggested by Baltes and Nesselroade (1979) for psychological research.

3. APC ANALYSIS OF MORTALITY: A BRIEF HISTORY

Although the impact on mortality trends of intergenerational variations in health was recognized in the 1920s, it was not until the 1970s that the cohort perspective was more generally adopted for the analysis of mortality trends. In general, cohort analysis and APC analysis has been motivated by one of two questions. The first focuses on the regularity of observed patterns and is associated with descriptive research; the second emphasizes the underlying mechanisms causing the regularity and is mainly associated with explanatory and epidemiological research (Hobcraft et al. (1985) make the same distinction; see also Hobcraft and Gilks, 1984). The two questions are the following:

a. Does an age distribution of a demographic phenomenon (e.g. mortality) exhibit a greater regularity when presented for a cohort than for a particular period? The age profile exhibited by period data confounds the effect of generational differences. This question is particularly relevant for demographic forecasting.
b. Do events and experiences early in life affect experiences later in life? In this perspective, cohort analysis derives its importance from the plausibility of biological mechanisms rather than from the use in forecasting.

According to Hobcraft et al. (1985, p. 103), Derrick (1927) was the first to argue that cohorts provided a more consistent basis for projecting mortality than did period rates. The conclusion was based on a graphical examination of the logarithms of age-specific death rates for England and Wales from 1841 to 1925, omitting the experience of World War I, which indicated that the ratio of mortality for one cohort to that of another cohort was approximately constant for all ages above 10. Caselli and Capocaccia (1989), however, review a study published in 1912 by Mortana, in which he studied the presence of possible selection effects in infancy on mortality at old ages. Pollard (1987, p. 58) lists studies which found that generation curves exhibit a greater degree of regularity. These studies are published in the 1920s and 1930s; in recent times, this regularity has not been observed to the same extent. Manton (n.d., p. 31) reports that period mortality schedules tend to overestimate cohort mortality rates. This is particularly so when part of the cohort is eradicated by a war. Since relatively healthy persons are selected for active service, they suffer great losses, while less healthy people are more likely to survive. This adverse selection leads to an overestimation of true mortality some decades later (Dinkel, 1985, p. 95). The selection is also in effect when mortality is studied by cause of death.

Explanatory research into the mechanisms underlying changes in mortality patterns focus on the impact of early experience on subsequent behavior. Kermack et al. (1934), studying time series of death rates of England, Scotland and Sweden, argued that the cohort differences in mortality was not a consequence of a series of independent conditions affecting successively older ages; instead, the health of a cohort was principally determined by environmental conditions encountered in its first 15 years of life. The authors also found that improvements in early childhood mortality followed mortality improvements in ages of maternity. They argued that early childhood mortality was closely linked to the health and physique of mothers. Kermack et al. adopted the life course perspective on cohort analysis long before it became popular in the 1980s when individual-level data became available. Preston and Van de Walle (1978), studying French data, and Caselli and Capocaccia (1989), using Italian data, demonstrated a positive relation between infant and child mortality and adult or old-age mortality (weakening effect). Others, however, stressed that high infant and child mortality result in lower mortality at higher ages because of a selection effect (e.g. Manton et al., 1981).

The introduction of cohort analysis in public health is generally attributed to Andvord (1921, 1930) and Frost (1939), who showed that apparent changes in age-specific rates of mortality from tuberculosis (TB) could be viewed as translations of declining TB mortality across cohorts with a relatively constant age profile of TB mortality. The authors believed that the TB infection occurred early in life and that the disease has a highly variable incubation period, tending to the lengthy (see Mason and Smith, 1985, p. 155). This implies that differences in infection rates in childhood largely determine differences in cohort experience. The authors suggested that, in the absence of effective chemotherapy, successive cohorts moved through life as though they had different probabilities of dying from TB assigned at birth. McKeown (1976), who has carried out one of the most authoritative research into causes of decline in mortality from micro-organisms, argues in the case of TB that changes in the probabilities of dying from TB are preceded and caused by improved nutrition (for a discussion, see Mason and Smith, 1985, pp. 156ff.). A major contribution of the study was the demonstration that the age distribution of mortality from TB was constant (regular) in cohorts rather than in periods and that period analysis may lead to erroneous conclusions.

Case (1956) adopted a cohort perspective in the study of long cancer in England and Wales for the period 1911-1954. The importance of cohort effects rested on the plausibility of biological mechanisms rather than on statistical tests. Case argued that the fact that successively younger cohorts were smoking cigarettes more heavily caused the cohort effects. Other references are listed in the bibliography and Appendix B.

A major research preoccupation of those European countries which were actively involved in World Wars I and II was examining the health and mortality situation at advanced ages of men who saw active service. The studies revealed two major findings. First, it has been shown in France, Italy and the Federal Republic of Germany that male cohorts which participated in World War I subsequently experienced higher mortality than adjacent cohorts who were not involved in the conflict. Second, in Italy and the Federal Republic of Germany, the same excess mortality has been detected among those who were born or were adolescent during the war years (Vallin, 1973, 1984 (France); Horiuchi, 1983 (Federal Republic of Germany); Caselli and Capocaccia, 1989 (Italy); Caselli et al., 1986 (Italy and France)). Boleslawski (1985) found a similar impact of World Wars I and II in Poland. In France, no notable weakening of the cohorts born during the World Wars was found (Wilmoth et al., 1988, p. 16). Anderson and Silver (1989) studied mortality data from the Soviet Union from 1958-59 to 1986-87 and found that males and females who were born during World War II and males who were adolescent during that time experienced significantly higher mortality as they aged than would have been expected on the basis of their age at a given time and the overall mortality conditions of the given period. The prolonged mortality effect on those who were adolescent during the war is attributed to the lasting effect of malnutrition on cardiovascular development (Horiuchi, 1983).

Caselli (1990) reports a remarkable observation for Italy. Very high levels of excess male mortality are found in the late 1960s for cohorts born during or just before World War I. She speculates that better living conditions allow more individuals to survive and make them more resistant to death until around age 50 (Caselli, 1990, pp. 239 and 245). In the Soviet Union, the rise in mortality of males in the working ages in the 1960s was attributed to World War II (Benyi and other Soviet scholars, quoted in Anderson and Silver, 1989, p. 477). Dinkel (1985) also suggested that the increase in male mortality in the 1960s might be attributed to the weakening effects of World War II. Anderson and Silver are reluctant for such an interpretation of the cohort mortality estimates because they can trace at most only 30 years of the mortality experience of any cohort (Anderson and Silver, 1989, p. 492).

4. STATISTICAL THEORY

The statistical theory of APC models is of a recent date. According to Hobcraft et al. (1985), the first properly identified APC model was specified by Greenberg et al. (1950). The age effects were parameterized through a beta distribution. The first author to make the linear identification constraint explicit was Beard (1963). Examples of APC analysis of mortality trends include Barrett (1973, 1980), Osmond and Gardner (1982), Osmond et al. (1982), Tu and Chuang (1983), Geddes et al. (1985), Mason and Smith (1985), etc. The state-of-the-art in the mid-eighties of the statistical theory of the APC model was discussed by several authors in the book edited by Mason and Fienberg (1985).

The application of an APC model to a time series of age-specific data raises a statistical problem, which is known as the identification problem and which received much attention in the literature. When the data are presented in an age-period table, as is common in APC studies, the cohort cannot unambiguously be identified. For instance, a 20-year old person who experiences an event in 1991, is born in 1970 or 1971. If the event occurs before the birthday, the person is born in 1970. The person belongs to the 1971 birth cohort, however, if the event occurs after the birthday. The cohort effect cannot uniquely be determined since the cohort is not properly measured. All that can be estimated is the difference between cohort effects. The problem is known as the identification problem. The identification problem is solved by equating two cohort effects or fixing a cohort effect to a given value (aliasing). Analogously, a person born in 1971 experiencing an event in 1991, may be either 19 years of age (if the event occurs before the birthday) or 20 years (if the event occurs after the birthday). If the data are arranged by year of occurrence of the event and year of birth, the age effect cannot be fully disentangled. The reason is not the linear relationship between age, period and cohort, as is suggested in most of the literature, but the inadequate measurement of age, period and cohort (see Willekens and Baydar, 1986; Robertson and Boyle, 1986; Osmond and Gardner, 1989). The measurement problem may be demonstrated graphically with the Lexis diagram (see Appendix A). The identification problem may be removed by

a. proper measurement of the timing of the event (date of occurrence, date of birth and age),
b. combining ages, cohorts or periods such that the number of effects to be determined reduces compared to the number of observations,
c. imposing restrictions on the values of the parameters (identification specifications),
d. substituting the age, period and/or cohort variables by other (better) proxies of life cycle stage, contemporary factors and historical factors, respectively.

The first approach was used by Willekens and Baydar (1986) and Robertson and Boyle (1986). The second approach is adopted in this paper. The third approach is followed in much of the traditional APC analysis (for a review, see Willekens and Baydar, 1986). The fourth approach is applied by Heckman and Robb (1985) and Blossfeld (1986) among others. The fourth approach is to be preferred if data permit.

In this paper, the APC model is presented as a special case of a generalized linear model (GLM). A similar approach was adopted by Willekens and Baydar (1986). The number of deaths is a random variable associated with a stochastic process. Model fitting consists of three interrelated steps, following McCullagh and Nelder (1983): (i) model selection (model specification or identification), (ii) parameter estimation, and (iii) prediction.

A. Model Selection

The model relates the outcome of the random process to the parameters of the process. The outcome is the number of events (deaths) in a particular interval, or any function of number of events. In this paper, we study the trend in death rates, defined as the ratio of the numbers of deaths and population at risk. The number and types of parameters are determined by the type of data that are available. One parameter is associated with each age, cohort and period.

B. Estimation

Given the model, we have to estimate the parameters from the data and obtain some measure of the accuracy with which we have estimated them.

C. Prediction

Prediction is concerned with the outcome of the actual random variable. Prediction is commonly thought of in the context of forecasting a future value of a variable. However, prediction is wider in scope and is used to indicate that the value assigned to a random variable is to be determined.

4.1. Model Selection

Models that we select to represent the data belong to the family of generalized linear models. An important characteristic of GLM's is that they assume independent observations. In case of non-independence, the variances will be larger than in the case of independent observations. It is assumed that deaths are generated by a Poisson process, hence the observed numbers of deaths follow a Poisson distribution. The Poisson assumption is justified when the death rate is low. In that case, the Poisson distribution is an adequate approximation of the binomial distribution, which describes binary response data (e.g. deaths/survivors) (McCullagh and Nelder, 1983, p. 74). The assumption that the number of deaths is an outcome of a Poisson process, has become widely accepted in the literature and is implicit in the log-linear analysis of mortality rates (see e.g. Holford, 1980; Laird and Olivier, 1981; Frome, 1983, with a discussion by Nelder, 1984; Egidi et al., 1990).

The dependent variable is the death rate, which is the ratio of the number of deaths and the total duration during which the population is exposed to the risk of dying. Since the exposure varies with the death rate, both the numerator and the denominator of the death rate are random variables and are interdependent. The dependence complicates the analysis substantially. Therefore, it is generally assumed that the denominator is fixed, i.e. independent of the number of deaths. If the death rate is small, the assumption is realistic. For a discussion of the issue, see Hoem (1984, pp. 41ff.) and Breslow and Day (1985, p. 57).

A major problem in model selection is the choice of variables to be included in the systematic part of the model. The strategy adopted in this paper is to associate one parameter with each age, period and cohort category.

Let n_{xtc} denote the observed numbers of deaths of age x, period t and cohort c. Let N_{xtc} denote independent random variables having Poisson distribution with positive parameter λ_{xtc}. λ_{xtc} is the product of the death rate and the duration of exposure to the risk of dying in year t by individuals of age x and cohort c, which is assumed to be fixed (L_{xtc}). The true value consists of two components: a systematic component, predicted by the model to be specified, and a random component. To be precise, the random component must be separated into two parts. One is a part due to our ignorance, i.e. the absence of a complete observation; the other part is due to the fact that the outcome of any random process is inherently uncertain even if we have all the necessary data to predict the outcome. No distinction between the two parts is made in this paper.

Let λ_{xtc} denote the systematic component and ε_{xtc} the random component. The model is:

$$n_{xtc} = \lambda_{xtc} + \varepsilon_{xtc} \qquad (1)$$

with $E(n_{xtc}) = \lambda_{xtc}$
$E(\varepsilon_{xtc}) = 0.$

A. The Systematic Component

The parameter λ_{xtc} of the Poisson distribution and λ_{xtc} are assumed to satisfy a model that is loglinear in a set Θ of unknown parameters. One parameter is associated with each of the ages, cohorts and periods. The systematic component is

$$\lambda_{xtc} = L_{xct}\, \alpha_x\, \beta_t\, \tau_c \qquad (2)$$

where $\Theta = \{\alpha_x, \beta_t, \tau_c\}$ and L_{xct} is the duration of exposure assumed to be given. Model (2) is the multiplicative formulation of the log-linear model. The additive formulation is obtained by taking the natural logarithm of both sides. In that case, the ln of the dependent variable is linear in the parameters.

The unknown parameters must be determined from the data. That can be shown by (i) writing the probability density of the outcomes of N_{xtc}, which gives the probability of observing any of the possible values of N_{xtc}, n_{xtc} say, given the model and data, and (ii) maximizing that probability. The maximum likelihood estimation of the parameters will be discussed after the presentation of the random component.

B. The Random Component

The independence and Poisson assumptions imply that the random variable N follows a Poisson distribution and that the probability of exactly n_{xtc} deaths in year t of persons of age x and cohort c, is given by the probability density function

$$Pr(N_{xtc}=n_{xtc}) = \exp[-\lambda_{xtc}]\lambda_{xtc}^{n_{xtc}}/n_{xtc}!$$

(3)

The Poisson distribution (3) is a member of the family of exponential probability density functions (McCullagh and Nelder, 1983). To show this, we rewrite (3) as follows

$$\Pr(N_{xtc}=n_{xtc}) = \exp[n_{xtc}\ln\lambda_{xtc} - \lambda_{xtc} - \ln n_{xtc}!] \quad (4)$$

Since the Poisson distribution is a member of the exponential family and the logarithmic transformation of the systematic component is linear in the parameters Θ, it is possible to estimate the parameters of the distribution by maximizing the likelihood of the parameters with respect to the observations on the random variable. We now proceed with the estimation.

4.2. Parameter Estimation

The parameters Θ are estimated by maximizing the likelihood of the outcomes of the independent Poisson processes, given the model (2) and the data. Since the logarithm is a monotonous increasing function, maximization of the log-likelihood is equivalent to maximization of the original likelihood. For a single observation n_{xtc}, the contribution to the likelihood is $n_{xtc}\ln\lambda_{xtc} - \lambda_{xtc}$. The log-likelihood of a set of observed flows n_{xtc}, where each flow is the outcome of a Poisson process with parameter λ_{xtc}, is:

$$L = \sum_{xtc}[n_{xtc}\ln\lambda_{xtc} - \lambda_{xtc} - \ln n_{xtc}!] \quad (5)$$

The maximization is not affected by the last term of (5), which may therefore be omitted.

If the model would perfectly predict the outcome of N_{xtc}, i.e. the maximum likelihood estimates are equal to the observations themselves ($\lambda_{xtc} = n_{xtc}$ and $\varepsilon_{xtc} = 0$), the likelihood is the maximum achievable, which is generally finite. To evaluate the goodness of fit of the model, we compare the likelihood achieved by the current model to the maximum of the likelihood achievable (i.e. the likelihood achieved by the full model). The logarithm of the ratio is known as the scaled deviance (see e.g. McCullagh and Nelder, 1983, pp. 24-25; GLIM Manual). The deviance is proportional to twice the difference between the log likelihoods:

$$S(n,\lambda) = -2\ln[L(\lambda,n)/L(n,n)]$$

$$= 2[\ln L(n,n) - \ln L(\lambda,n)]$$

Large values of S indicate low values of $L(\lambda,n)$ relative to the full model, increasing lack of fit. For the Poisson distribution, the deviance is

$$S(n,\lambda) = 2 \sum_{xtc} [n_{xtc} \ln (n_{xtc}/\lambda_{xtc}) - (n_{xtc} - \lambda_{xtc})] \quad (6)$$

If a constant term ϕ, which is known as the nuisance parameter, is included in the model it is generally the case that $\Sigma(n_{xtc} - \lambda_{xtc}) = 0$ so that

$$D(n,\lambda) = S(n,\lambda)\, \phi$$

may be written in the more usual form of the log-likelihood ratio which is often used as a test in the analysis of contingency tables

$$D(n,\lambda) = 2 \sum_{xtc} n_{xtc} \ln (n_{xtc} / \lambda_{xtc}) \quad (7)$$

In order to determine the unknown Θ parameters with maximum likelihood, we need to maximize the log-likelihood function with respect to the parameters. This results in a set of normal equations which need to be solved for the unknown parameters. The GLIM package, which uses generalized weighted least square, was applied. The weights are inversely related to the variances of the estimates. The algorithm uses the Fisher's scoring method. If the model is log-linear, the scoring method and the Newton-Raphson method reduce to the same algorithm (McCullagh and Nelder, 1983, p. 33; Aitken et al., 1989, pp. 324ff).

4.3. Prediction

The most probable number of deaths that are consistent with the available data and the model are given by the expected values of the N_{xtc}, which is λ_{xtc}. The expected death rate may be written as follows:

$$\lambda_{xtc}/L_{xtc} = \kappa\, \alpha_x\, \beta_t\, \tau_c / L_{xtc} \quad (8)$$

where the parameters are restricted

$\alpha_1 = 1$, $\beta_t = 1$ and $\tau_c = 1$.

Alternative restrictions may be used.

5. APPLICATION

5.1. The Anderson-Silver Analysis

Anderson and Silver (1989) studied the trends in Soviet adult mortality since 1958-59 and outlined three competing explanations for the trends: cohort effects, period effects and data quality changes. They were the first to apply age-period-cohort analysis to Soviet mortality data. The authors used reported five year age-specific death rates (ASDR) for ages 5 to 59. For the USSR as a whole, age-specific rates are published for the years 1958-59 through 1986-87 in two-year intervals. For republics, the published ASDR's start with 1970-71 for two-year periods until 1986-87. The USSR data cover a period in which two major reversals in adult mortality occurred: a reversal from a declining mortality to an increasing mortality in 1964-65 and a reversal to a declining mortality in 1985.

The mortality rates that are available every two years for age groups of five years must be mapped onto birth cohorts used in the analysis. The authors took the central year of birth of a particular age group in a given two-year period and assigned the value for that group in that year to the five-year birth cohort in which the central birth year fell. For example, for persons aged 10-14 in 1970-71, the central birth year was 1958 (1970.5-10-2.5). The ASDR of that group in that year was assigned to the birth cohort of 1955-59 (Anderson and Silver, 1989, pp. 487ff). The result is an age-period table with several data points (observations) for most cohorts. Because the number of period observations exceeds the number of period effects to be estimated, the identification problem does not arise. The data are given every two years, the period effects are estimated for five-year periods (i.e. the effects for the two-year periods within a five-year period are assumed to be equal). The cohorts considered are from 1900-1904 to 1975-79. Note that three observations are available for the 1975-79 cohorts (mortality rate at ages 5-9 in 1982-83, 1984-85 and 1986-87).

Anderson and Silver identify the age, period and cohort effects in a two-step procedure. They first estimate the age and period effects exhibited by the age-period table of logarithm of death rates. The residuals are used to determine the cohort effects, which were generated by forcing the regression through the origin. OLS regression was used to obtain the effects. The analysis of Anderson and Silver was repeated for the USSR as a whole using exactly the same data and the same method (OLS) but a different package (GLIM). Identical parameter estimates are obtained; they are shown in Table 1 in the columns labelled 'Anderson&Silver'. The results may be compared with Tables 2 and 5 in the Anderson-Silver (AS) article. In comparing the figures, the following should be kept in mind. First, the effects shown by Anderson and Silver are the natural logarithm of the effects shown in Table 1. Second, the period effects in both Table 1 and the AS paper are normalized using the 1970 values (1970 = 1 in the multiplicative model and 0 in the additive formulation of the log-linear model). Third, the cohort effects are scaled differently in the AS analysis and our analysis. In the AS analysis, the cohort effects are normalized such that the effect of the 1900-1905 cohort is equal to the main effect (-0.0859 for males in the additive formulation and 0.91767 in the multiplicative formulation of the model). We scaled the effects using the effect of the 1920-24 cohort as the unit. Fourth, AS show an effect parameter for the 1975-79 cohort, whereas we do not. We exclude the parameter from the table (not from the analysis) because the effect of the 1975-79 cohort in the maximum likelihood estimation was aliased (see below).

5.2. Our Analysis

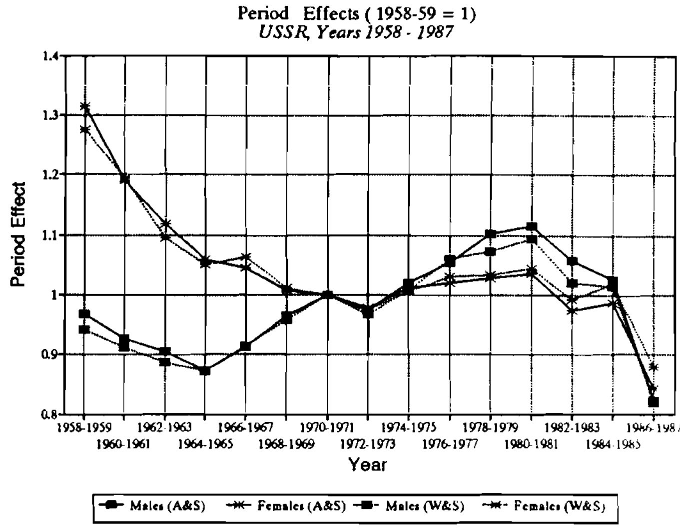
The method for estimating the APC model presented in Section 4 differs from the OLS estimation procedure used by Anderson and Silver in two ways:

- Anderson and Silver infer the cohort effects from the residuals of the age and period effects while we determine the age, period and cohort effects simultaneously.

Table 1. Comparison of the Anderson and Silver analysis.

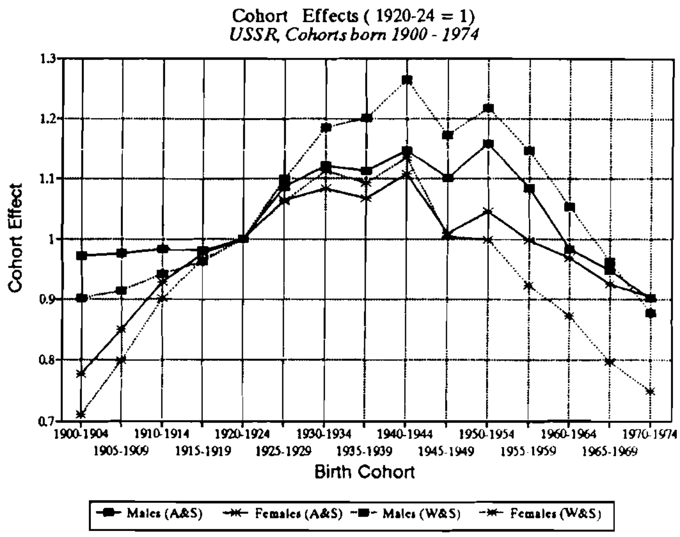
Period Effects (1958-59 = 1)

	Anderson&Silver		Willekens&Scherbov	
Year	Males	Females	Males	Females
1958-1959	0.968	1.315	0.941	1.276
1960-1961	0.926	1.190	0.911	1.193
1962-1963	0.905	1.118	0.887	1.095
1964-1965	0.872	1.059	0.871	1.050
1966-1967	0.912	1.045	0.914	1.062
1968-1969	0.964	1.008	0.956	1.011
1970-1971	1.000	1.000	1.000	1.000
1972-1973	0.974	0.978	0.967	0.975
1974-1975	1.019	1.011	1.007	1.005
1976-1977	1.054	1.019	1.060	1.031
1978-1979	1.101	1.028	1.073	1.033
1980-1981	1.115	1.035	1.093	1.044
1982-1983	1.058	0.974	1.020	0.993
1984-1985	1.025	0.986	1.015	1.018
1986-1987	0.822	0.842	0.818	0.879



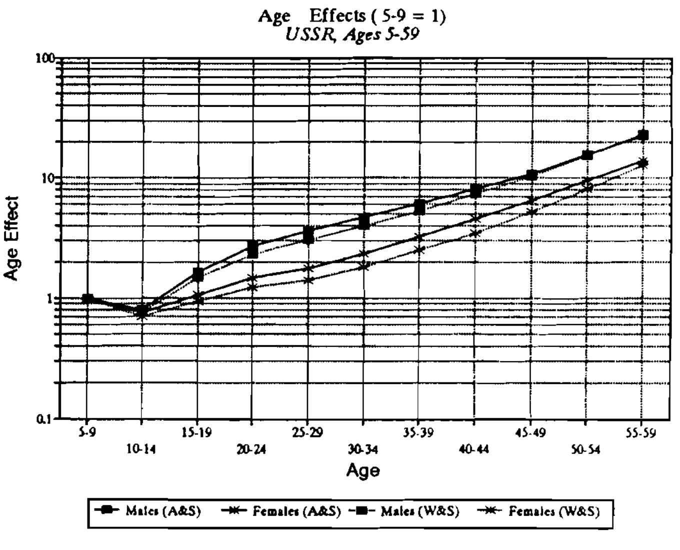
Cohort Effects (1920-24 = 1)

	Anderson&Silver		Willekens&Scherbov	
Birth cohort	Males	Females	Males	Females
1900-1904	0.972	0.777	0.903	0.710
1905-1909	0.976	0.851	0.915	0.799
1910-1914	0.983	0.927	0.942	0.901
1915-1919	0.981	0.976	0.963	0.967
1920-1924	1.000	1.000	1.000	1.000
1925-1929	1.086	1.065	1.099	1.062
1930-1934	1.121	1.083	1.185	1.113
1935-1939	1.112	1.067	1.200	1.092
1940-1944	1.147	1.108	1.264	1.135
1945-1949	1.101	1.009	1.171	1.005
1950-1954	1.158	1.045	1.217	0.998
1955-1959	1.084	0.997	1.147	0.923
1960-1964	0.983	0.968	1.054	0.873
1965-1969	0.948	0.925	0.963	0.798
1970-1974	0.902	0.905	0.878	0.750



Age Effects (5-9 = 1)

	Anderson&Silver		Willekens&Scherbov	
Age	Males	Females	Males	Females
5-9	1.000	1.000	1.000	1.000
10-14	0.797	0.749	0.761	0.702
15-19	1.627	1.063	1.458	0.933
20-24	2.690	1.469	2.315	1.215
25-29	3.618	1.749	3.068	1.394
30-34	4.734	2.335	4.003	1.800
35-39	6.073	3.228	5.328	2.498
40-44	8.105	4.510	7.316	3.490
45-49	10.894	6.476	10.329	5.212
50-54	15.701	9.656	15.534	8.158
55-59	22.511	14.176	23.313	12.871



- Anderson and Silver use the OLS procedure while we use the general weighted least square algorithm for exponential family regression models developed by Nelder and Wedderburn (1972) and implemented in GLIM to generate maximum likelihood estimates.

In order to measure the impact of differences in method, we fitted the APC model to exactly the same data as Anderson and Silver. The results are shown in Table 1 in the columns labelled 'Willekens&Scherbov' and in the associated graphs. The period and age effects are not much different. The period effects in our analysis are less pronounced in the extreme periods which is due to the high variances during these periods. The major difference is in the cohort effects, which might have been expected. The effect of weighting and the simultaneous estimation of age, period and cohort effect is that part of the cohort effects become more significant. Part of the effects that are attributed to period and age in the OLS are in fact due to cohort differences.

Since the completion of the Anderson-Silver paper, mortality data have become available for the years 1987, 1988 and 1989. The mortality rates are available annually since 1987. In order to estimate the age, period and cohort effects of mortality in the Soviet Union, we made use of the additional information. Consequently, the data used in the subsequent analysis consist of the data used by Anderson and Silver, augmented with data for the years 1987, 1988 and 1989. Note that our data include the year 1987 twice (1986-87 and 1987). In addition, we include the age groups 0-4, 60-64 and 65-69. The mapping of age groups onto birth cohorts is the same as described in the Anderson-Silver paper.

The APC analysis was carried out for males and females simultaneously. The model included the effect of age, period, cohort, sex and the interaction effect between age and sex. The model was fitted to mortality data for the USSR as a whole and for each republic. In order to compare the age, period and cohort effects exhibited by the data of the USSR with those exhibited by mortality in the USA, the APC analysis was carried out for males only.

A. Mortality of the USSR

The results for the USSR are shown in Figure 1. The period effects exhibit the increase in mortality since 1964-65 and the sharp drop following the drastic measures taken in May 1985 against drunkenness and alcoholism in the USSR (Andreev, 1990a, p. 15). The increase in mortality in 1987-89 is attributed to the economic crises (Andreev, 1990a, p. 18). The cohort effect shows a rapid decline since the second World War. The mortality of children who are either born during the war or grew up during the war is higher than that of other cohorts. The finding is consistent with that of Bednyi and other Soviet scholars who claim that World War II permanently affected the health and subsequent death rates of the young cohorts (quoted by Anderson and Silver, 1989, p. 477). A similar weakening effect (for males) was found in Poland (Boleslawski, 1985), Germany (Horiuchi, 1983; Dinkel, 1985), Italy and France (Caselli, 1990).

USSR

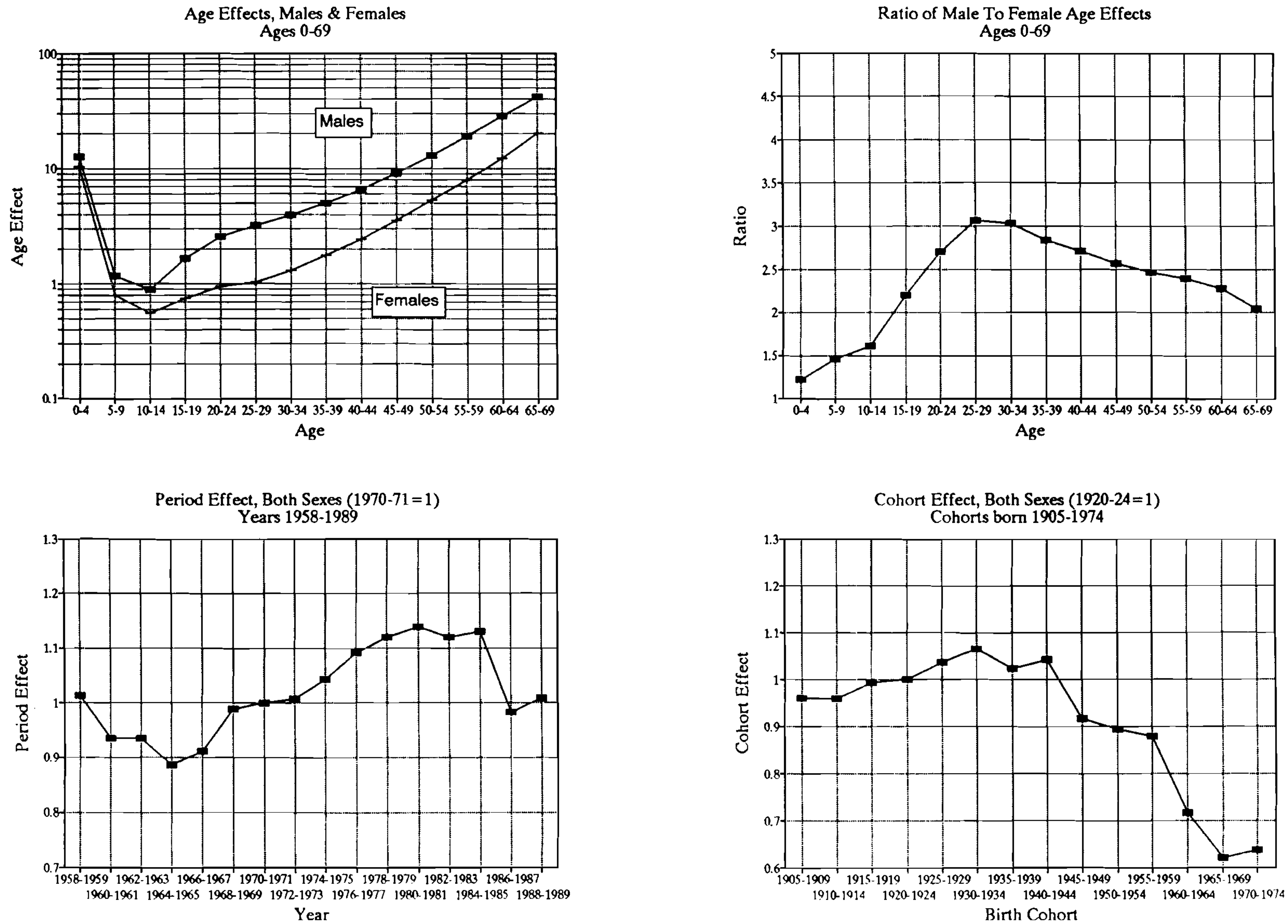


Figure 1. Results of APC analysis for the USSR.

Part of the rise in the period effects during the 1960s and the 1970s has been attributed to historical rather than contemporary factors. World Wars I and II would have a weakening effect on the cohorts who were born or were very young during the war. The weakening effect of World War I would become pronounced in the 1960s when the surviving children of the war pass age 60. Our present analysis does not enable us to completely separate the effect of the wars from the contemporary factors in the 1960s and 1970s. If worsening health conditions in the 1960s affect children of the war differently from other cohorts, the effect can only be captured by a period-cohort interaction. The addition of period-cohort interaction effects to the APC model increases the size of the design matrix substantially. We did not include the interaction term due to computer memory problems. Caselli (1990) follows a different approach to capture the period-cohort interaction effects measuring the long-term impact of the two World Wars. She includes in the APC model the mortality level between ages 0 and 15 years. The parameter associated with this variable measures the effect on mortality at later ages of the level of mortality during the first part of life. Based on the analysis of Italian data, she speculates that better living conditions (after the war) not only allow more individuals to survive (increased cohort effect), but also make survivors (of the war) more resistant to death until around age 50 (Caselli, 1990, p. 245).

The combined effects of period and cohort are clearly demonstrated in Figure 2. The three-dimensional figure shows the product of period and cohort effects for each year and cohort. The age effects are excluded from the data. The figure reveals that the period effects become pronounced in the mid-1960s and suppress the rise in cohort effects. That would indicate a significant independent effect of contemporary factors. The cohort effects are revealed by the shifts in period effects.

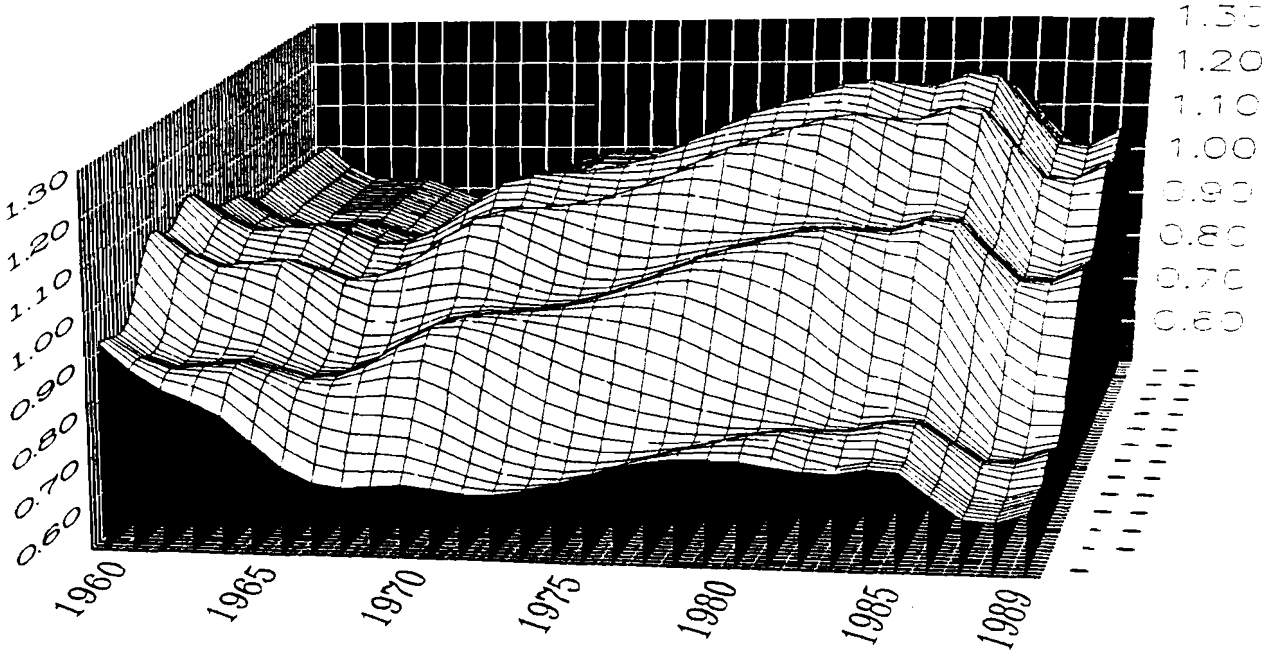


Figure 2. Cohort and period effects, USSR, male, 1958-1989.

B. Mortality of the Regions of the USSR

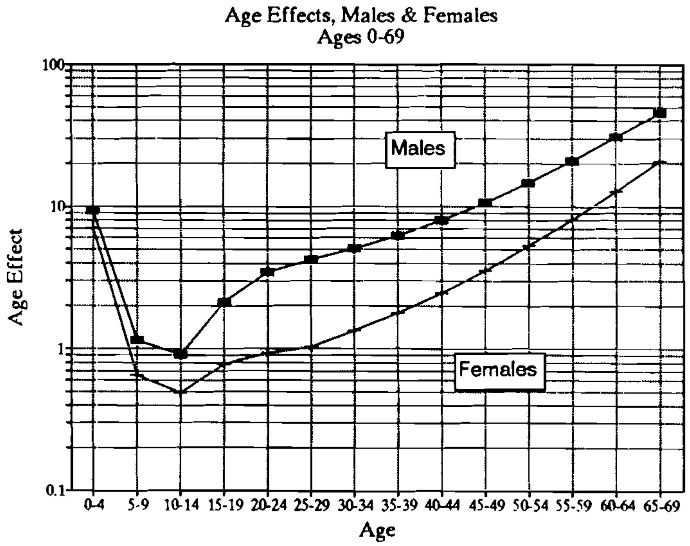
The regional study of age, period and cohort effects of mortality yield interesting differences (Figure 3a-o). The age effects of male and female mortality differ greatly in regions with a large proportion of European nationalities. Male mortality is substantial higher than female mortality at all ages except the lowest. The difference is particularly large at young adult ages. For instance, males of age 25-30 in Russia, Lithuania and Latvia have a mortality that is not less than 4 times that of females in the same age group. The difference is attributed to the high male probability of dying from accidents and injuries (Andreev, 1990b, p. 110). In the Asian regions, the differences between male and female adult mortality are limited. However, infant mortality is high.

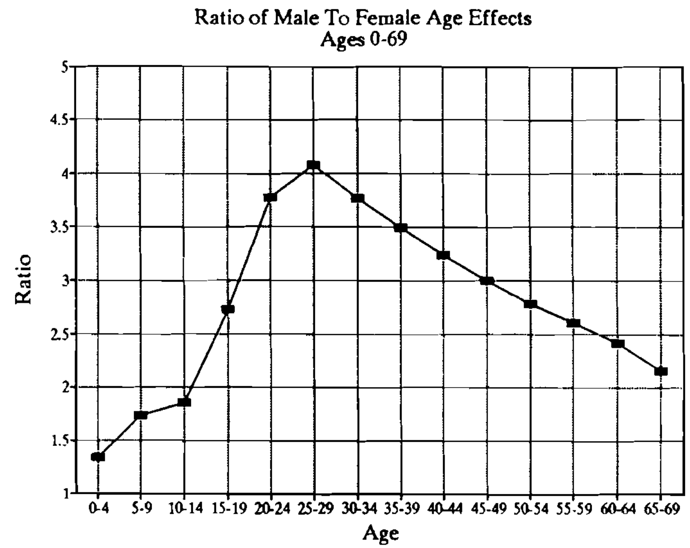
Figure 4a-b shows the ratio between the age effect parameters for males in each republic to the age effect in Russia. The figure is an outcome of an APC model that includes the age-region interaction. Adult male mortality is higher in Russia than in any other republic.

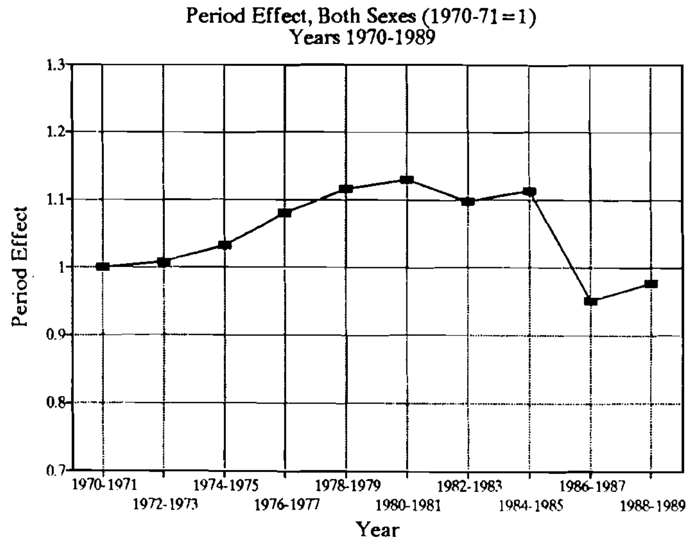
The drop in period effects of mortality after 1985 is particularly pronounced in the regions with European nationalities. In Azerbaijan, Turkmenia, Tajikistan, Georgia and Armenia, the period effect of the measures taken in May 1985 are absent. The earthquake in Armenia (1988) pushed the period effect to a very high level. The increased mortality of the cohorts born after the 1945-49, may in part be due to a confounding effect of the earthquake (period-cohort interaction). The mortality pattern in Armenia is most difficult to capture by an APC model that contains main effects only. The standard deviance for Armenia was 93.17, whereas the deviance is less than 10 for all other republics except Tajikistan.

The period effects of mortality are increasing in most regions since 1987 due to the economic crisis. The effect was limited or absent in Georgia, Tajikistan, Ukraine, Moldavia and Azerbaijan. In Tajikistan and Uzbekistan, however, the children born during 1970-74 have a much higher mortality than previous generations. This might indicate that young children are carrying the heaviest burden of the crisis in these regions.

Russia







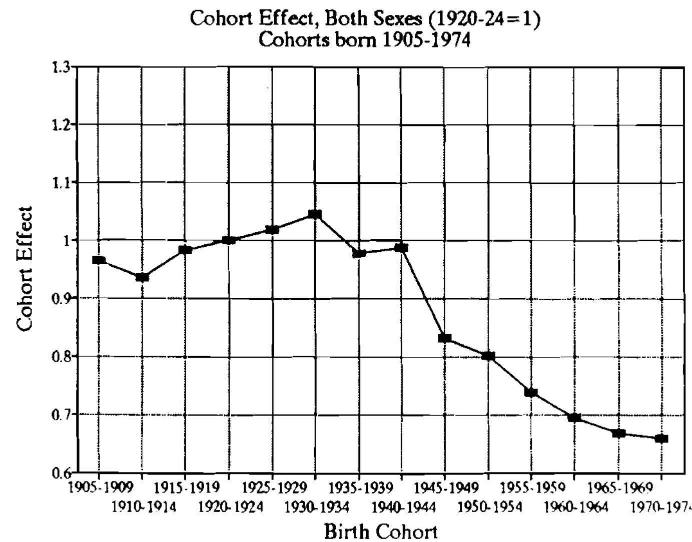
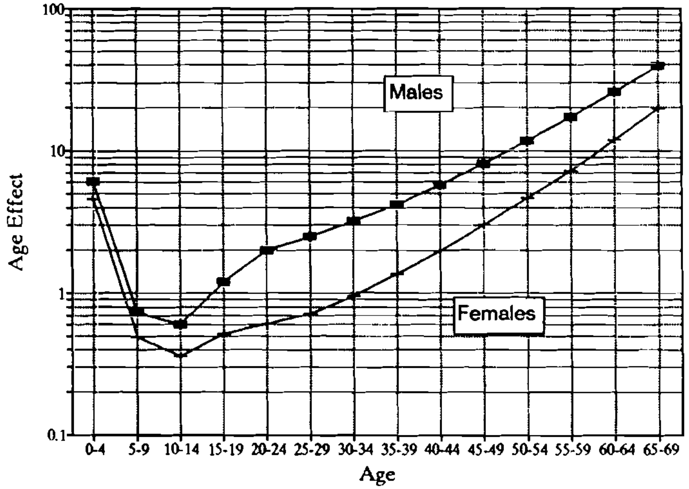


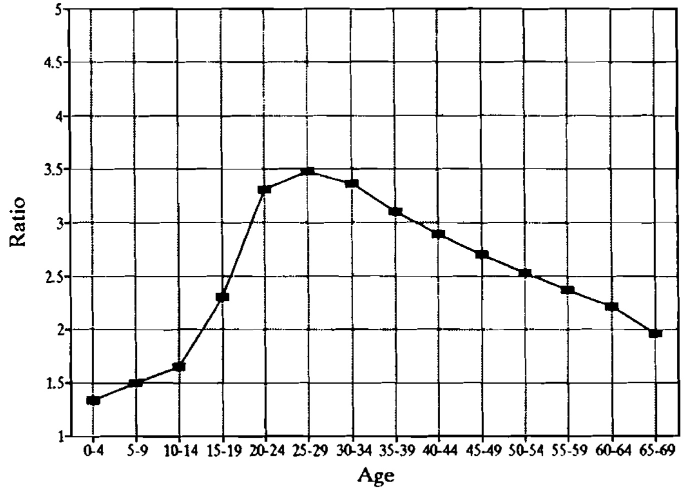
Figure 3a-o. Regional APC analyses.

Ukraine

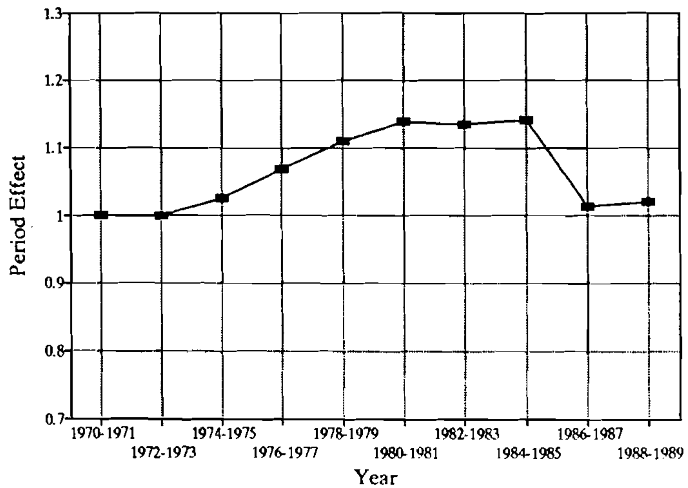
Age Effects, Males & Females
Ages 0-69



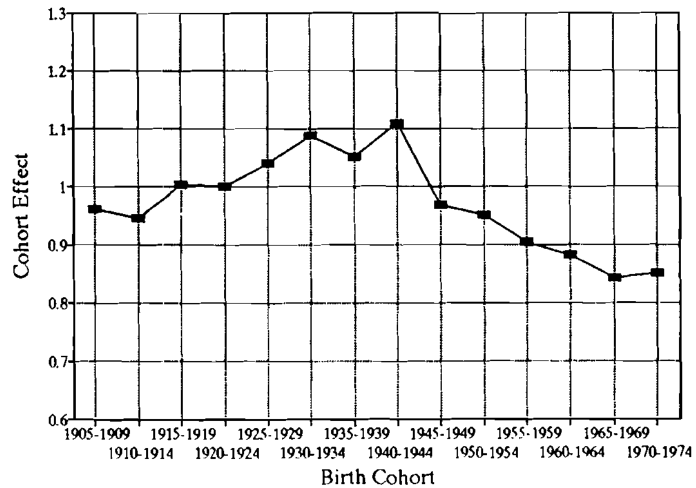
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

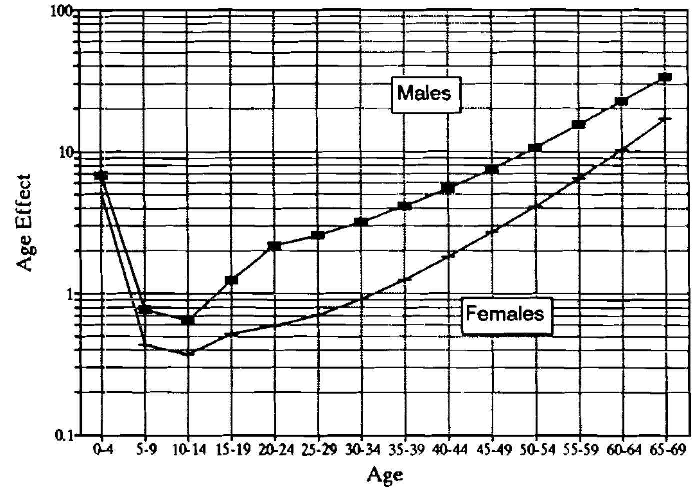


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

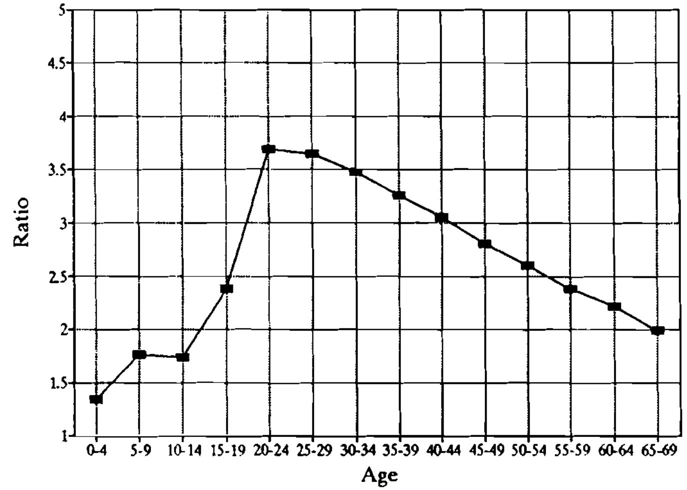


Byelorussia

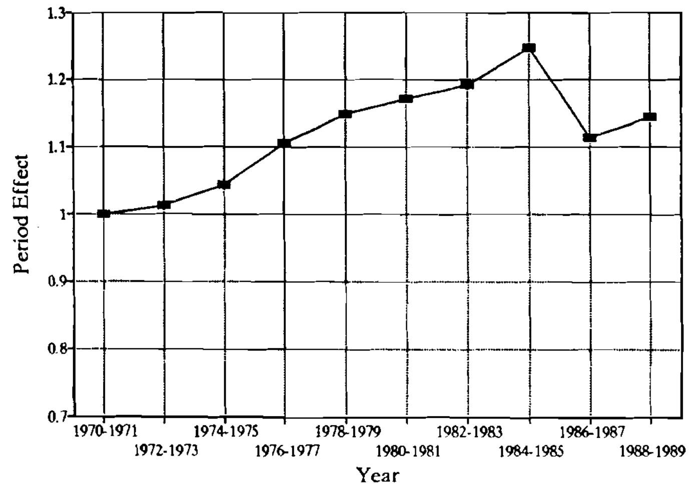
Age Effects, Males & Females
Ages 0-69



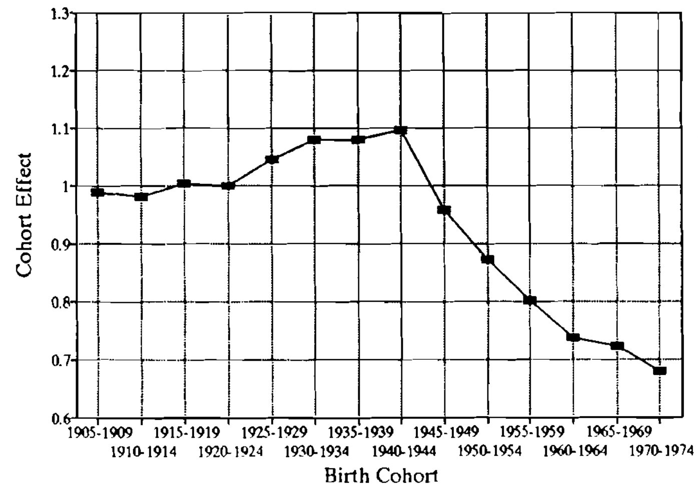
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71 = 1)
Years 1970-1989

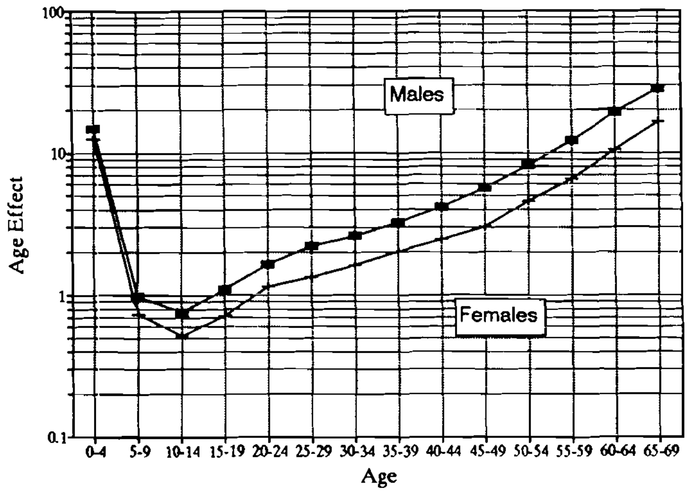


Cohort Effect, Both Sexes (1920-24 = 1)
Cohorts born 1905-1974

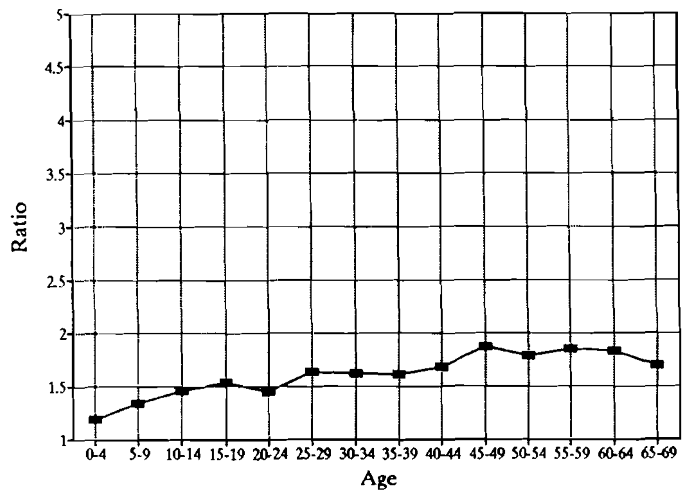


Uzbekistan

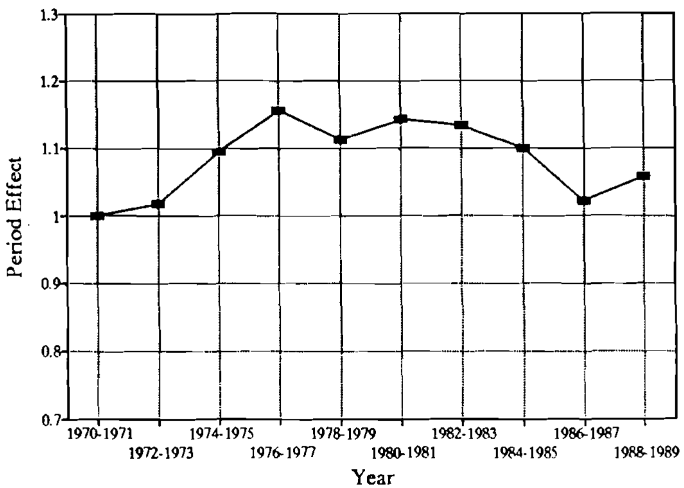
Age Effects, Males & Females
Ages 0-69



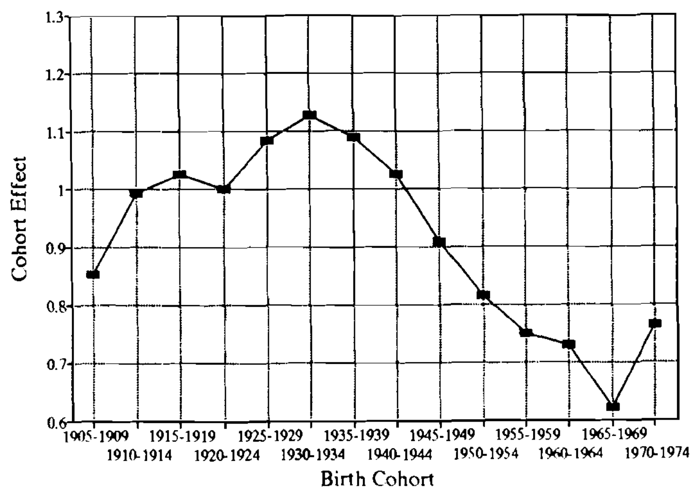
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71 = 1)
Years 1970-1989

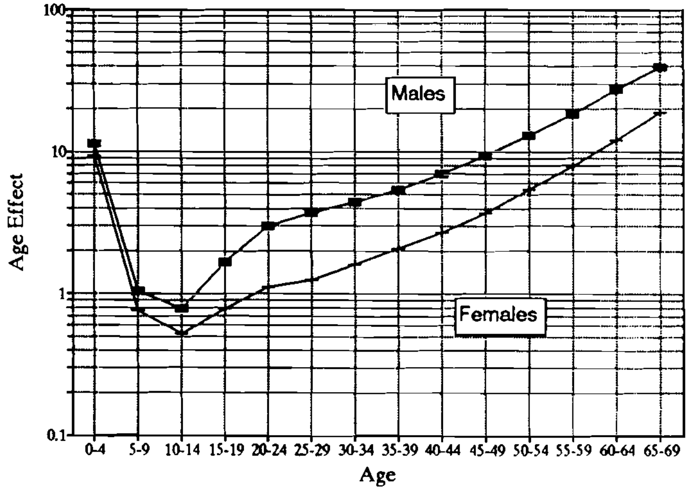


Cohort Effect, Both Sexes (1920-24 = 1)
Cohorts born 1905-1974

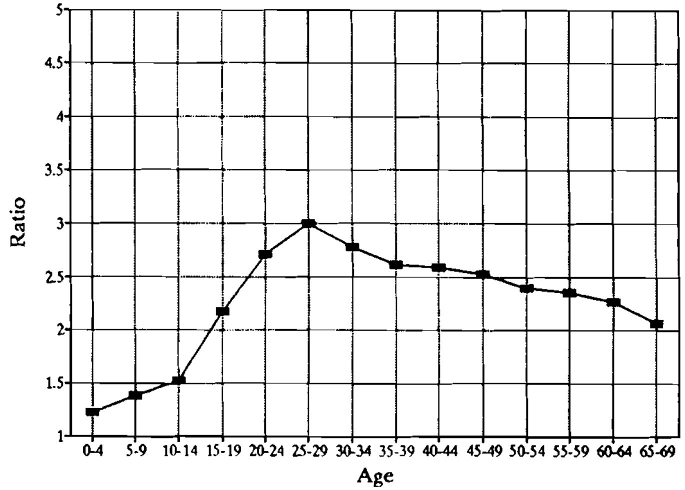


Kazakhstan

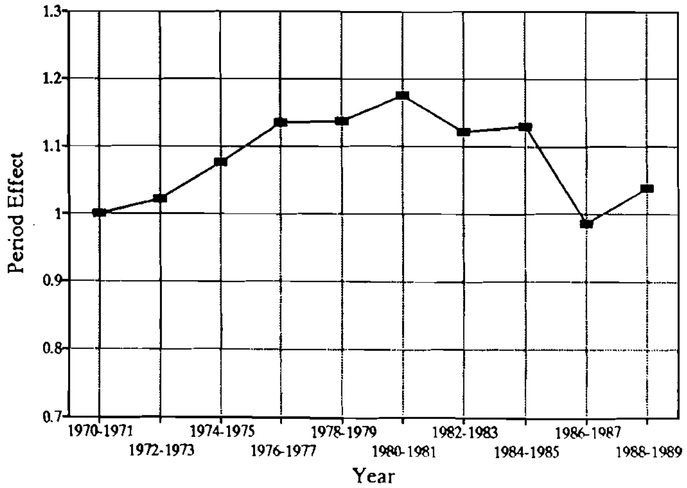
Age Effects, Males & Females
Ages 0-69



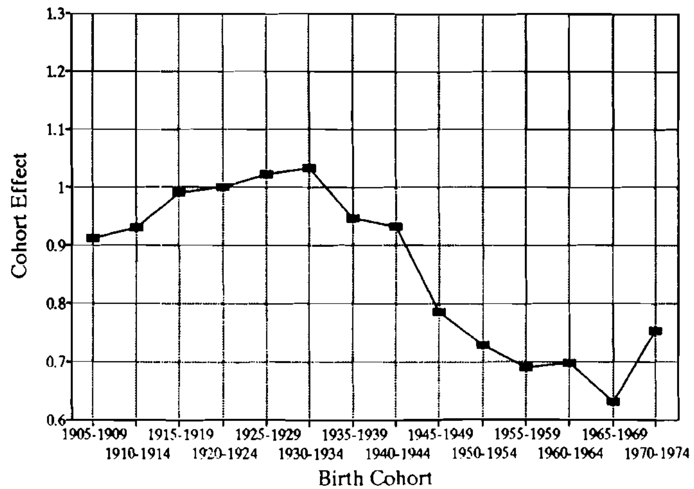
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

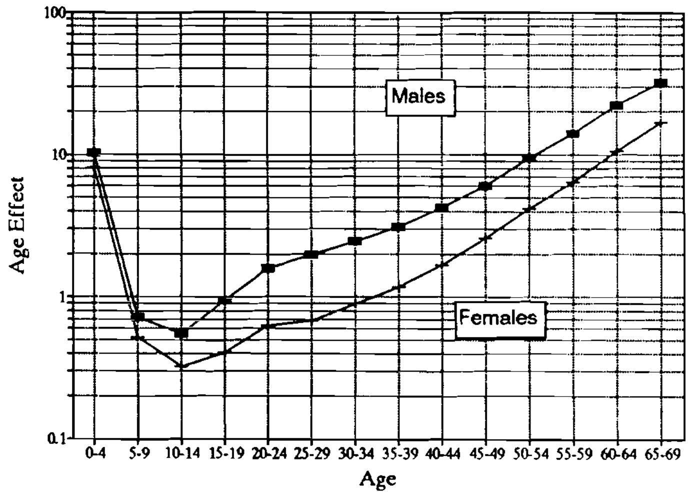


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

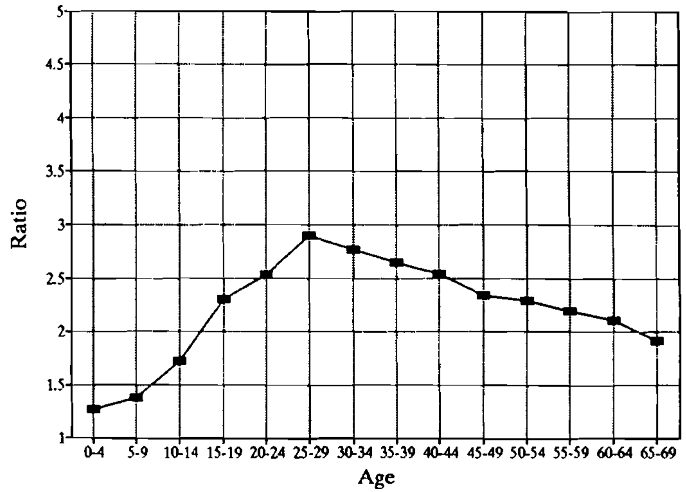


Georgia

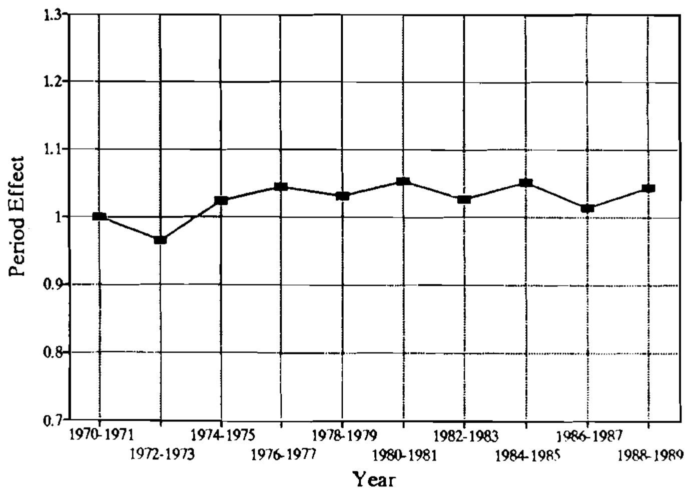
Age Effects, Males & Females
Ages 0-69



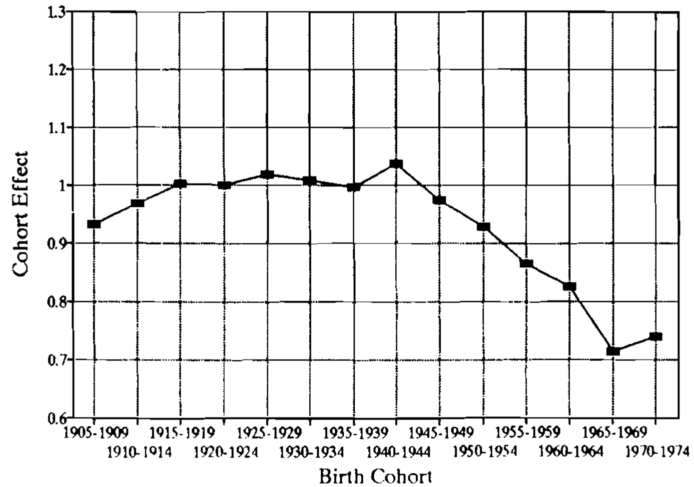
Ratio of Male To Female Age Effects
Ages 0-69



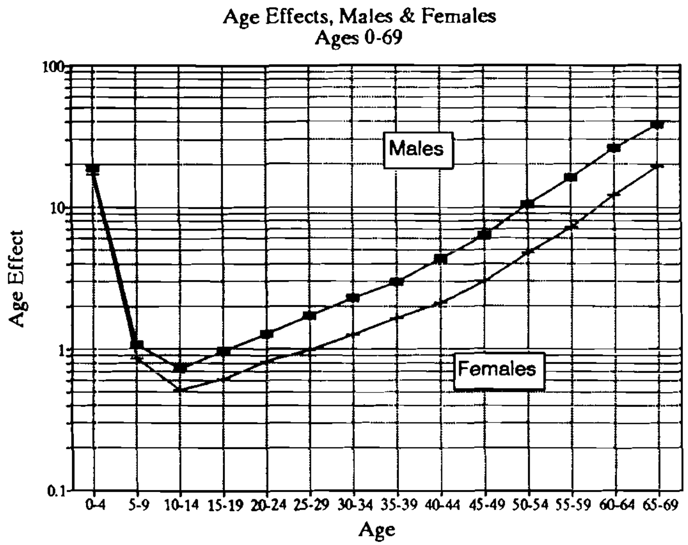
Period Effect, Both Sexes (1970-71=1)
Years 1970-1989



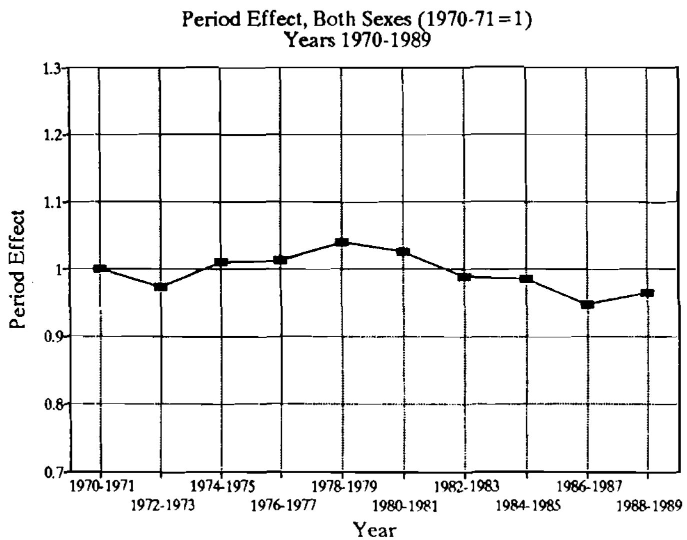
Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

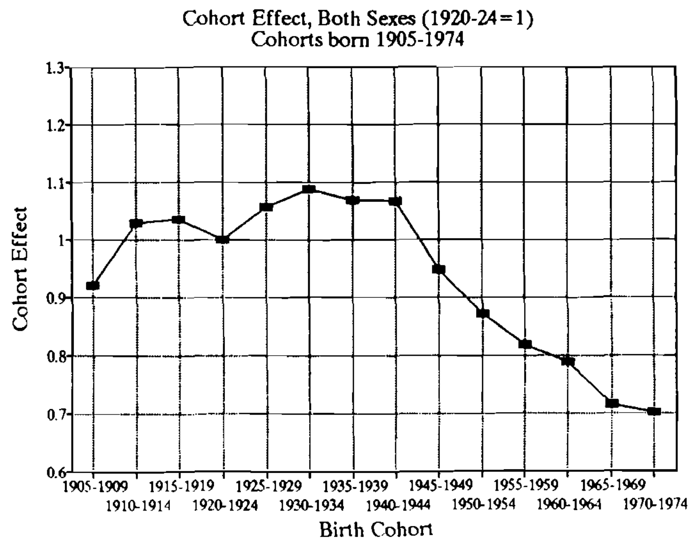


Azerbaijan


Age Effects, Males & Females
Ages 0-69
Age Effect
100
10
1
0.1
Males
Females
0-4 5-9 10-14 15-19 20-24 25-29 30-34 35-39 40-44 45-49 50-54 55-59 60-64 65-69
Age

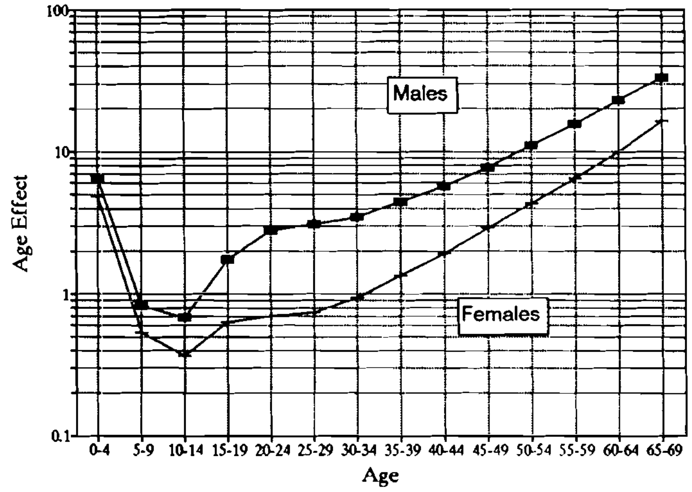

Ratio of Male To Female Age Effects
Ages 0-69
Ratio
5
4.5
4
3.5
3
2.5
2
1.5
1
0-4 5-9 10-14 15-19 20-24 25-29 30-34 35-39 40-44 45-49 50-54 55-59 60-64 65-69
Age


Period Effect, Both Sexes (1970-71=1)
Years 1970-1989
Period Effect
1.3
1.2
1.1
1
0.9
0.8
0.7
1970-1971 1974-1975 1978-1979 1982-1983 1986-1987
1972-1973 1976-1977 1980-1981 1984-1985 1988-1989
Year

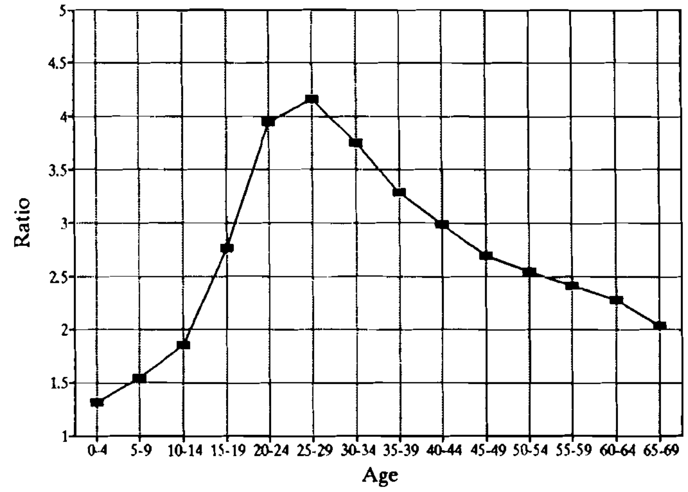

Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974
Cohort Effect
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
1905-1909 1915-1919 1925-1929 1935-1939 1945-1949 1955-1959 1965-1969
1910-1914 1920-1924 1930-1934 1940-1944 1950-1954 1960-1964 1970-1974
Birth Cohort

Lithuania

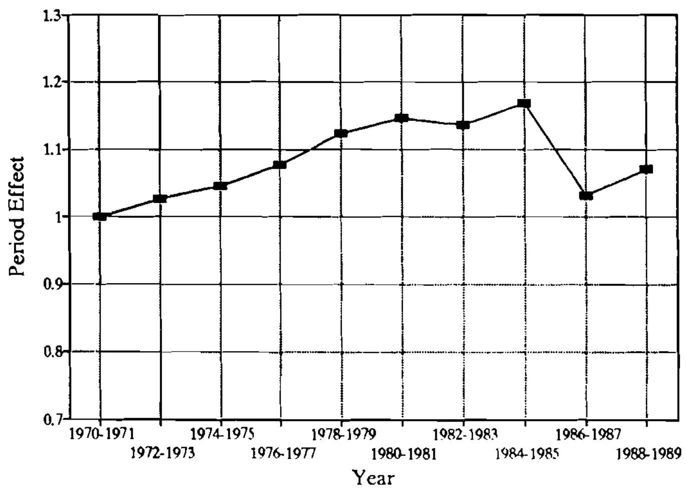
Age Effects, Males & Females
Ages 0-69



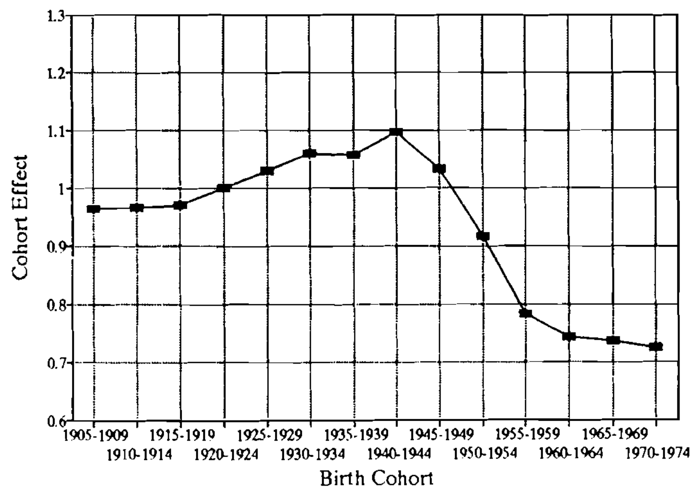
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

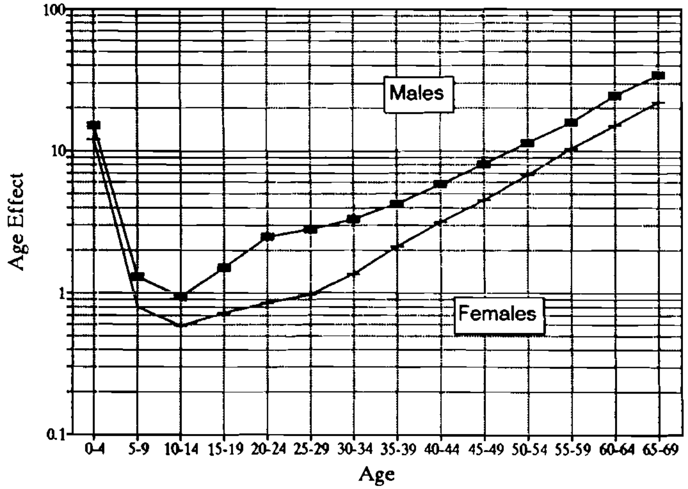


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

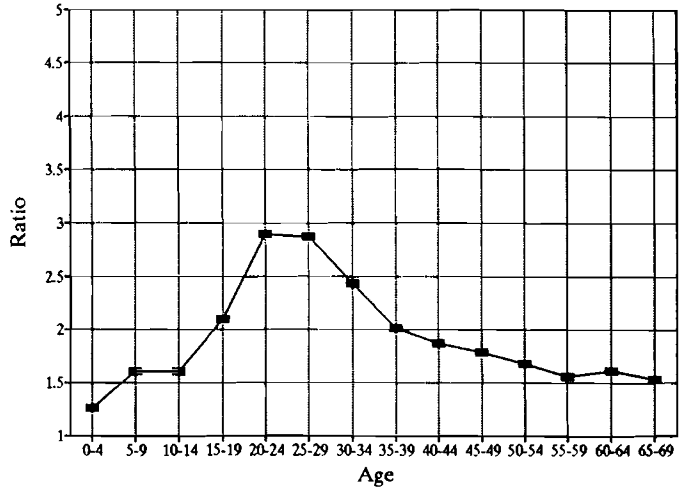


Moldavia

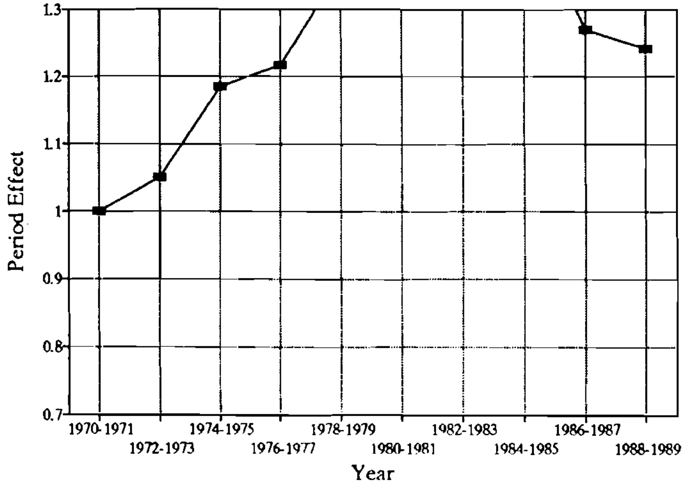
Age Effects, Males & Females
Ages 0-69



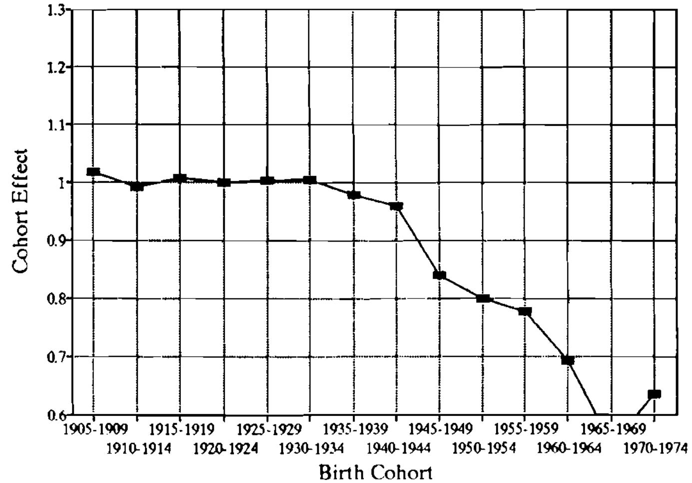
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

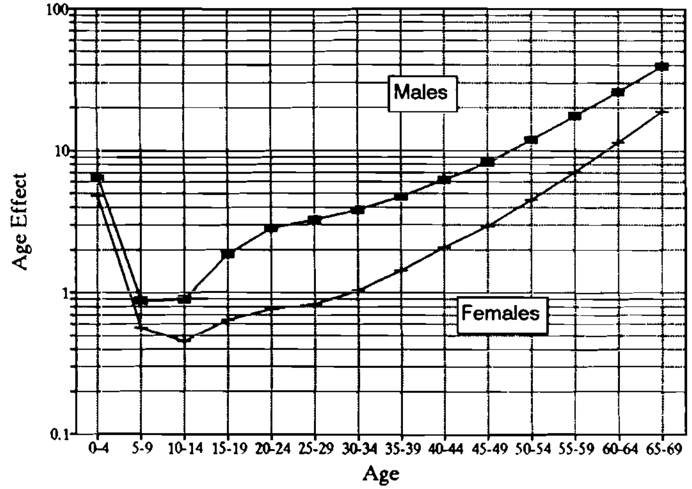


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974



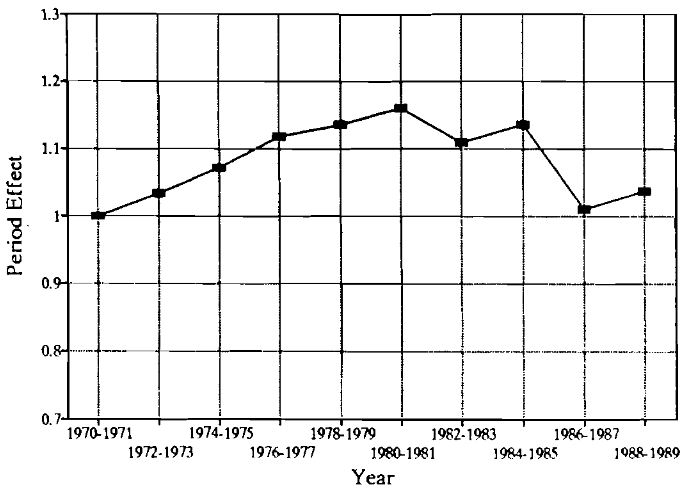
Latvia

Age Effects, Males & Females
Ages 0-69

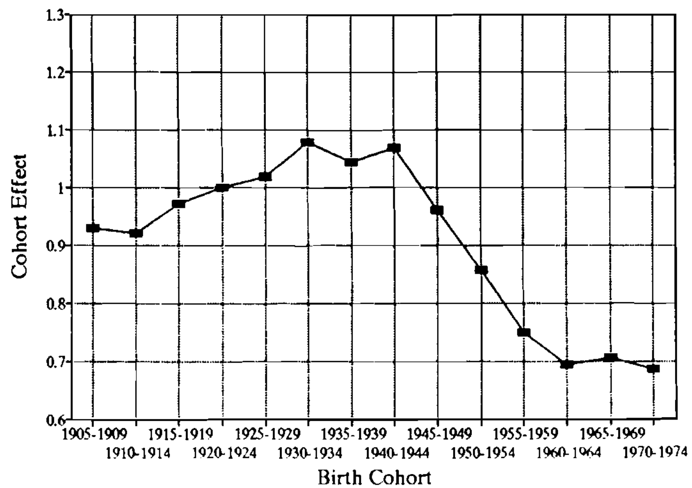


Ratio of Male To Female Age Effects
Ages 0-69

Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

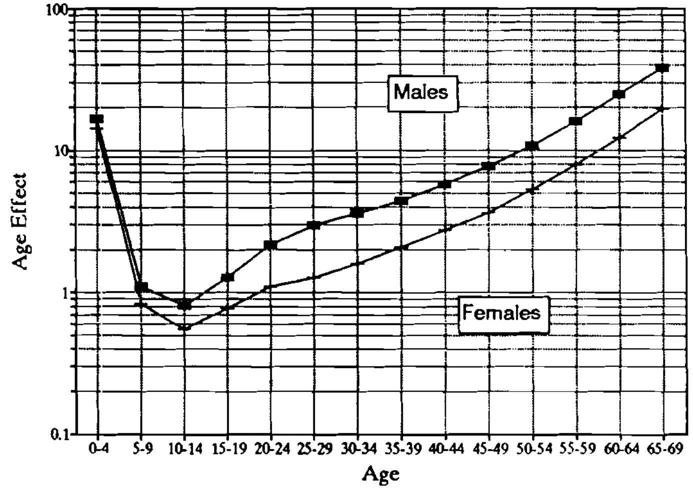


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

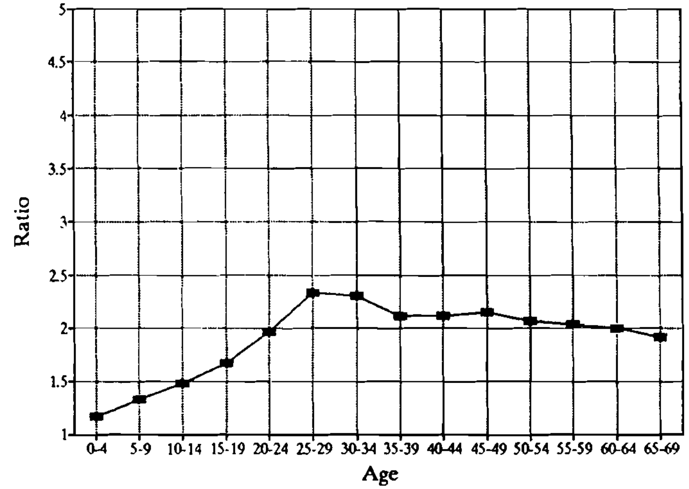


Kirghizia

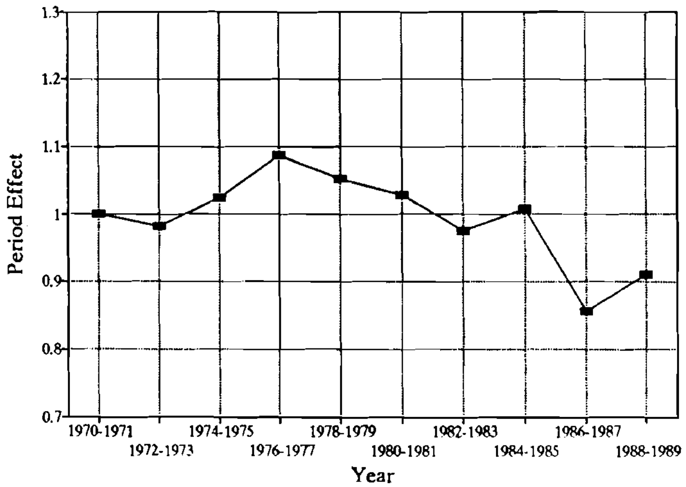
Age Effects, Males & Females
Ages 0-69



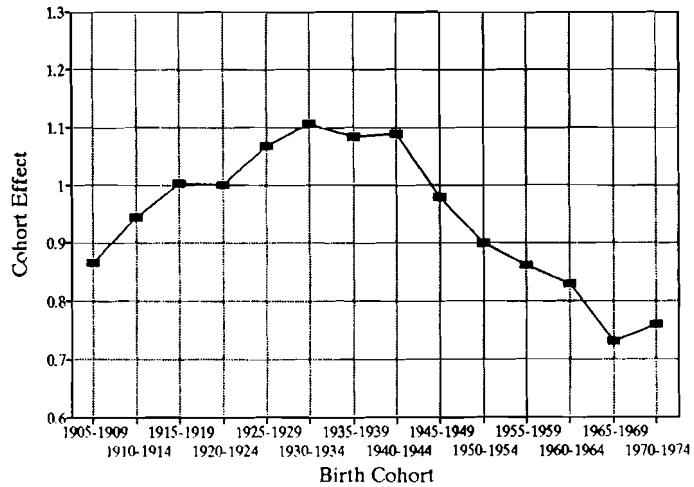
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

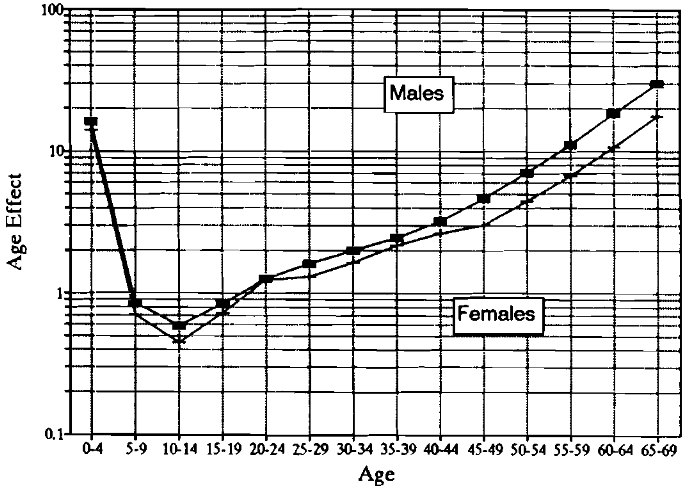


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

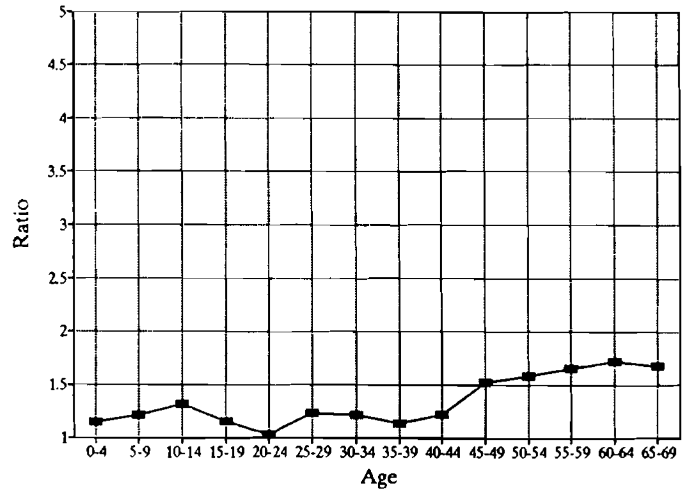


Tajikistan

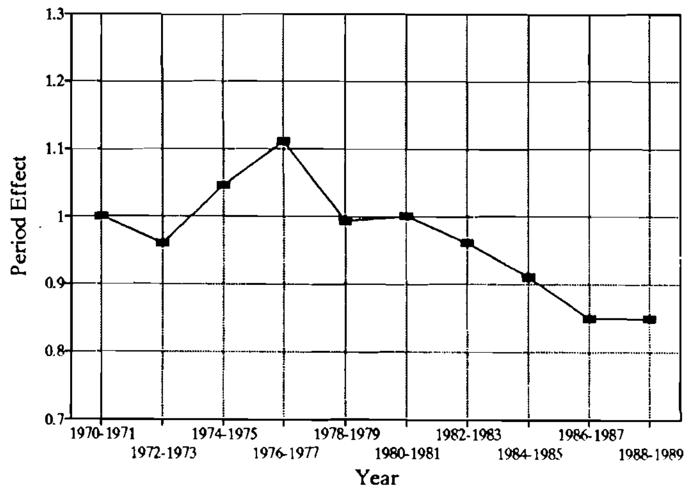
Age Effects, Males & Females
Ages 0-69



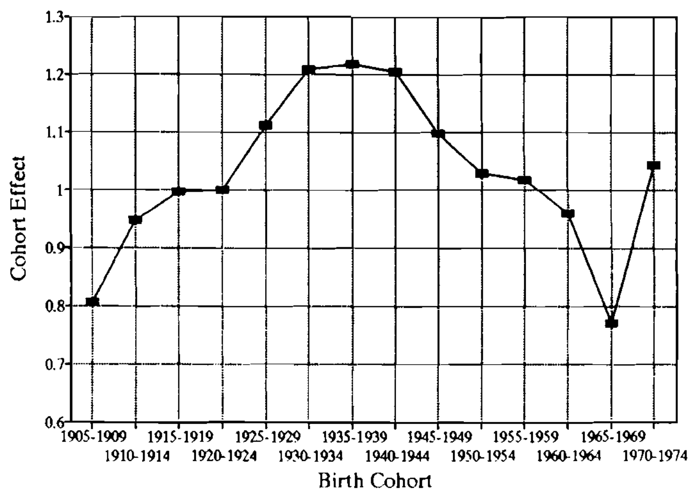
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71 = 1)
Years 1970-1989

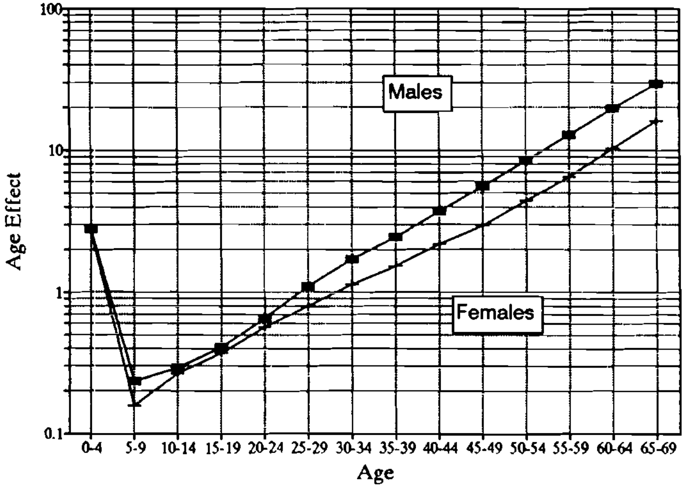


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

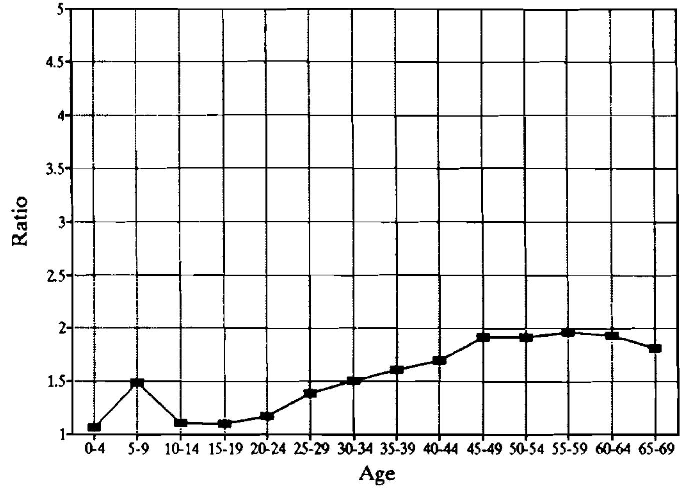


Armenia

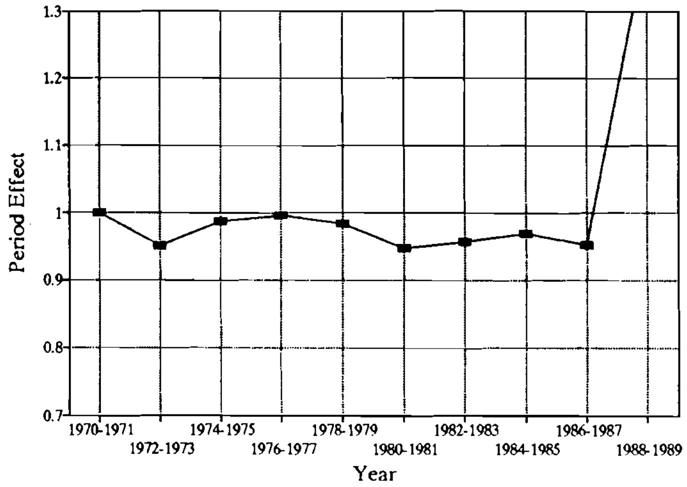
Age Effects, Males & Females
Ages 0-69



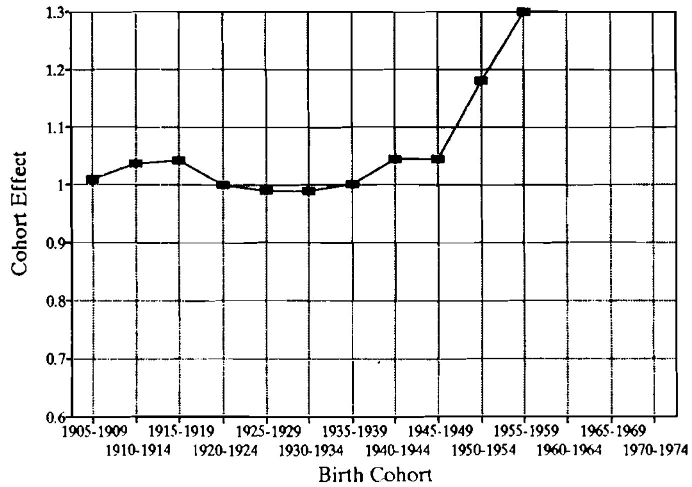
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

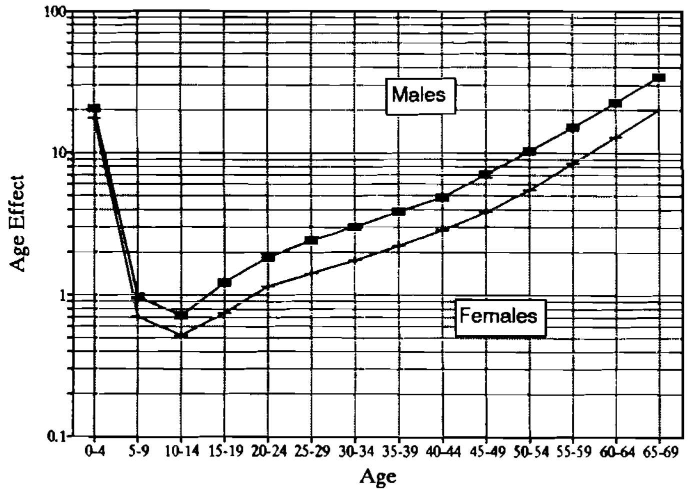


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

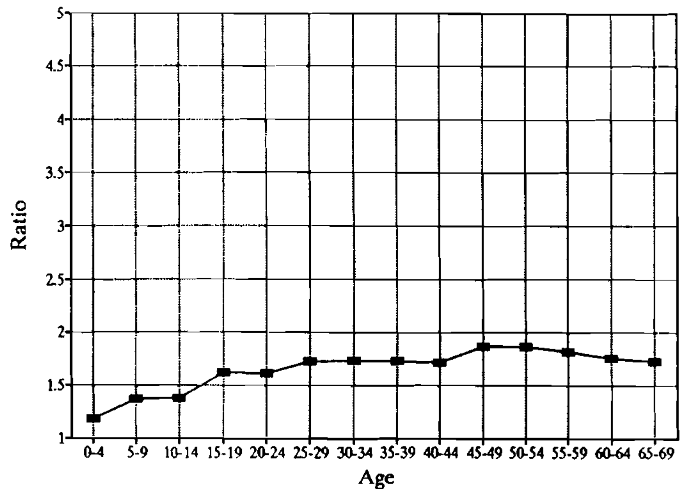


Turkmenia

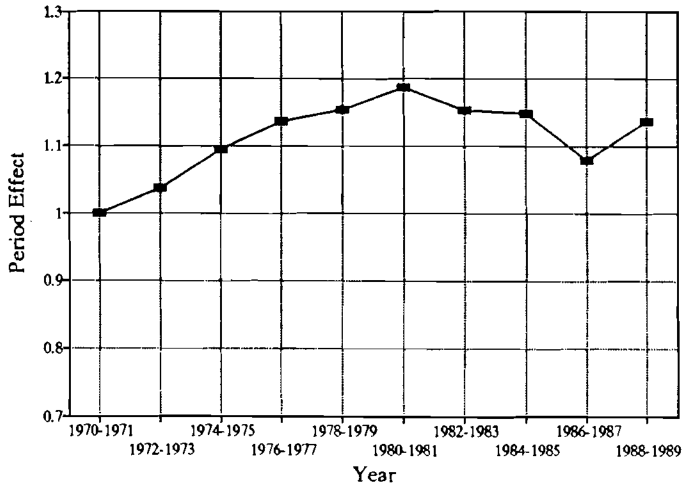
Age Effects, Males & Females
Ages 0-69



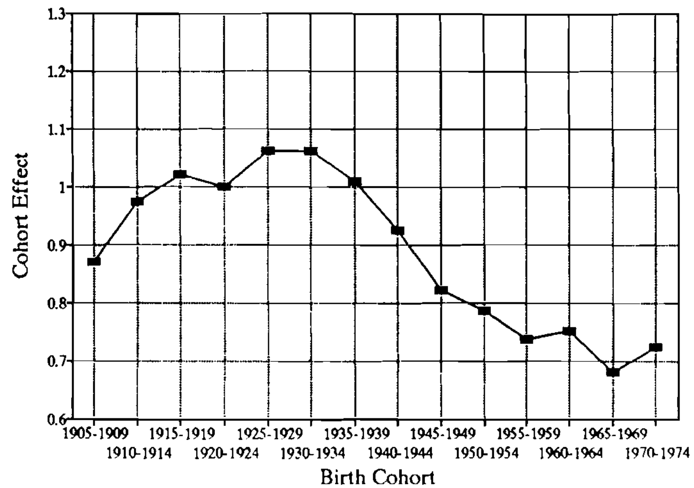
Ratio of Male To Female Age Effects
Ages 0-69



Period Effect, Both Sexes (1970-71=1)
Years 1970-1989

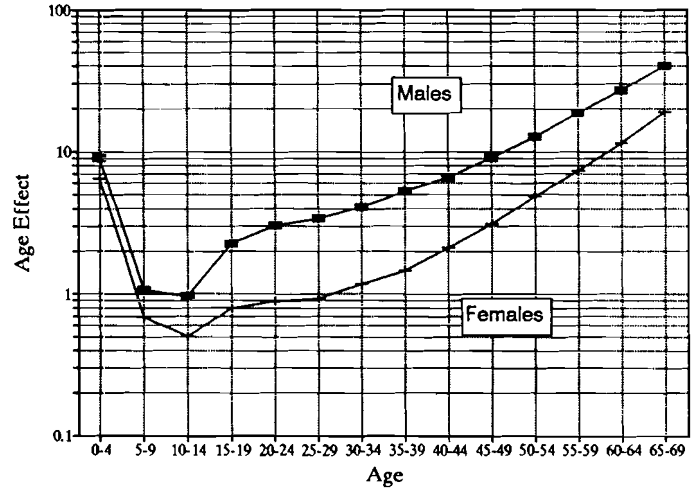


Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974

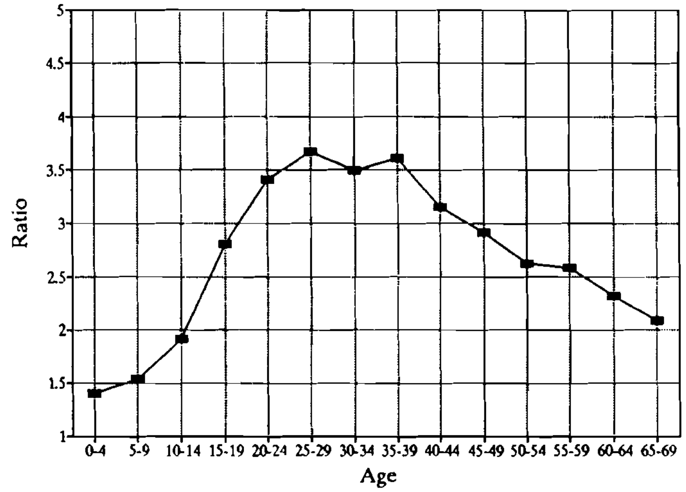


Estonia

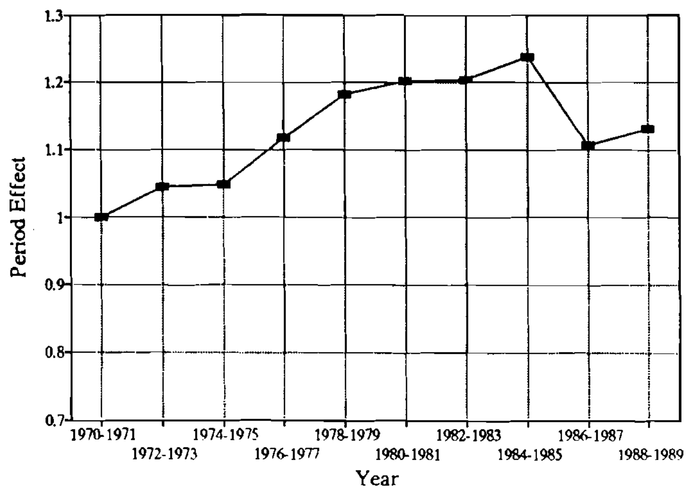
Age Effects, Males & Females
Ages 0-69



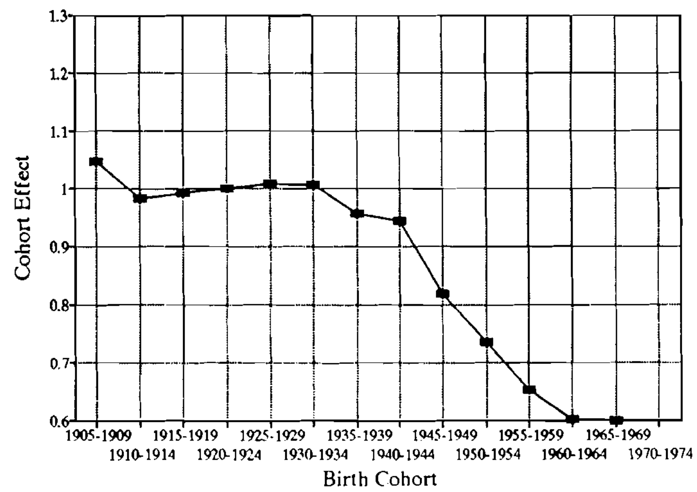
Ratio of Male To Female Age Effects
Ages 0-69

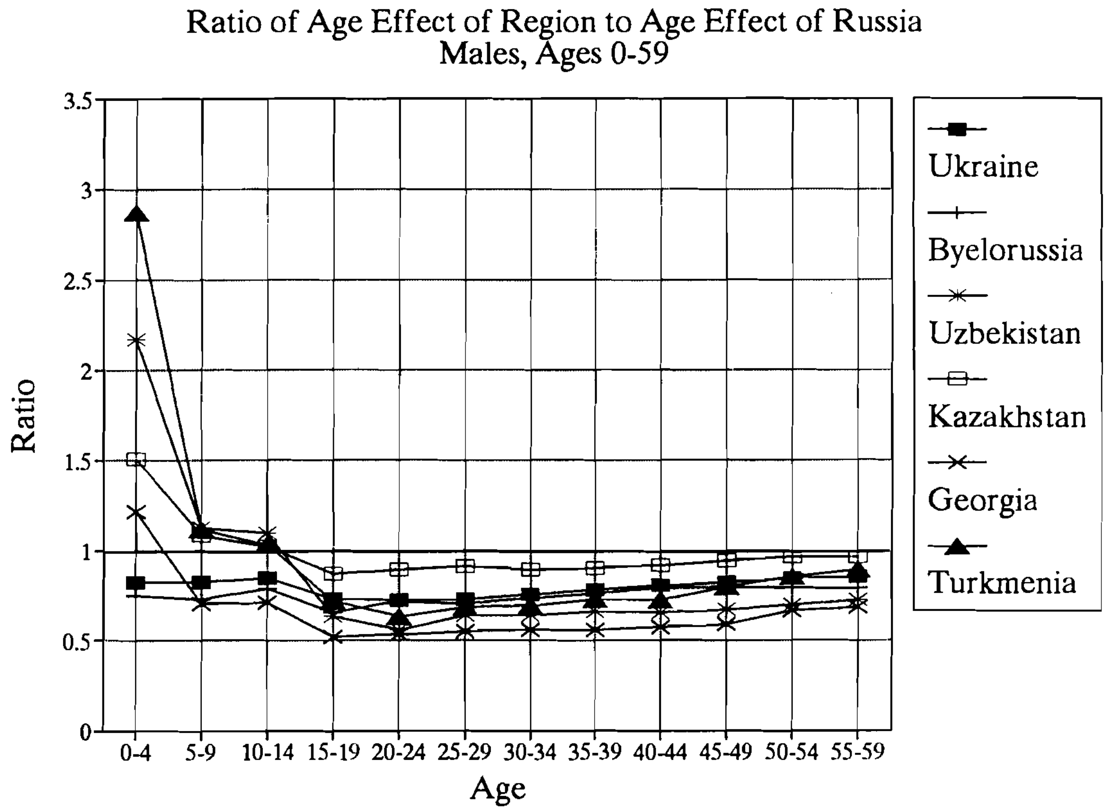


Period Effect, Both Sexes (1970-71=1)
Years 1970-1989



Cohort Effect, Both Sexes (1920-24=1)
Cohorts born 1905-1974





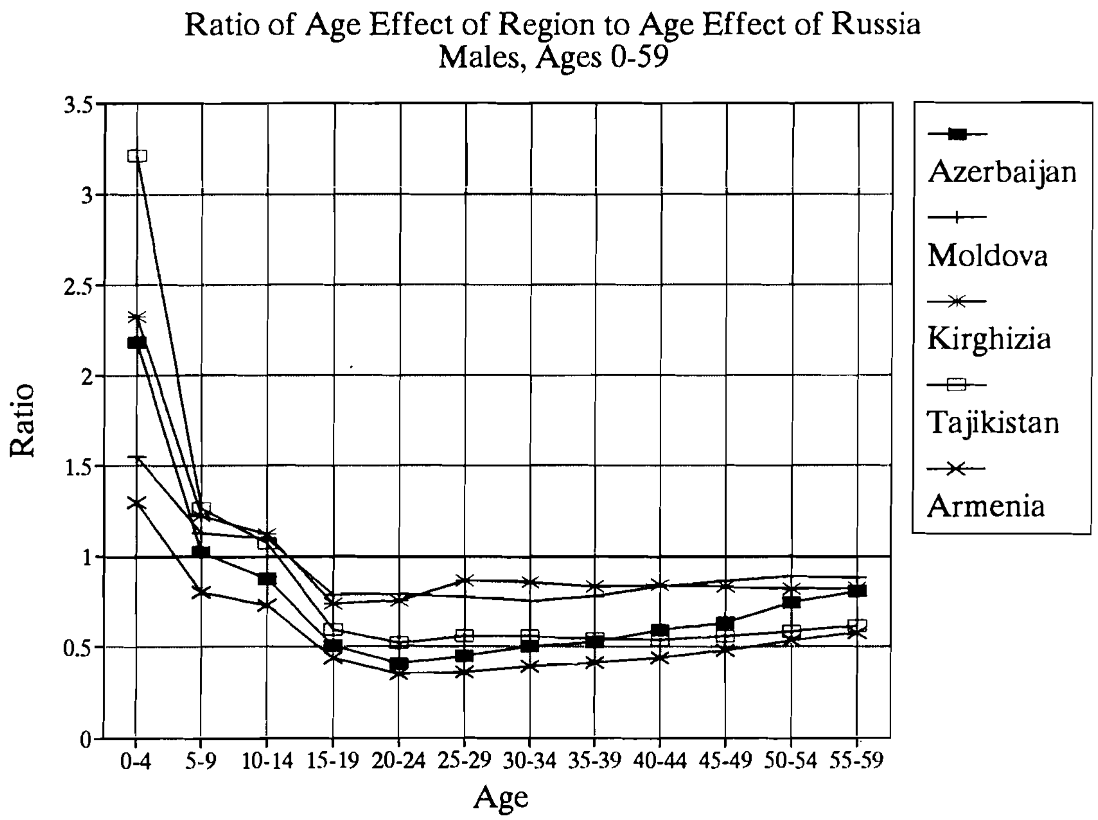


Figure 4a-b. Age effects of mortality in various regions divided by age effects in Russia.

C. Mortality of the USSR and the USA: A Brief Comparison

Since the mid-1960s, the mortality of the USSR is diverging substantially from mortality in the USA and other countries in the West. In order to compare the varying trends, an age-period-cohort model was estimated to US data on male mortality from 1948 to 1984. The data consist of annual, five-year mortality rates from ages 5-9 to 65-69. The estimation was done independently for the USSR and the USA.

The results are shown in Figure 5a-c. In the early sixties, the contemporary factors in the USSR were more beneficial to mortality than in the USA. The picture changed dramatically since the late 1960s: the contemporary factors in the USA pushed mortality down, whereas mortality increased in the USSR. The cohort and age effects to some extent neutralize the effects of period factors. Until the 1960 cohort, the cohort effect in the USSR and the USA developed in opposite directions. The increased mortality of cohorts born between 1940 and 1960 in the USA cannot yet be explained. The cohort mortality started to decline much earlier in the USSR than in the USA because the level of cohort mortality was very high for the cohorts born during or before World War II.

The age effects represent a classical case of cohort-inversion. The cohort-inversion model states that cohorts experiencing particularly hard or good times early in life will respond inversely later in life (Hobcraft et al., 1985, p. 93). In the USA, the age effects are lower than in the USSR up to age 45 and are higher after that age. The difference is particularly large in age group 30-35. By way of comparison, the period age-specific mortality rates of 1964 are also shown.

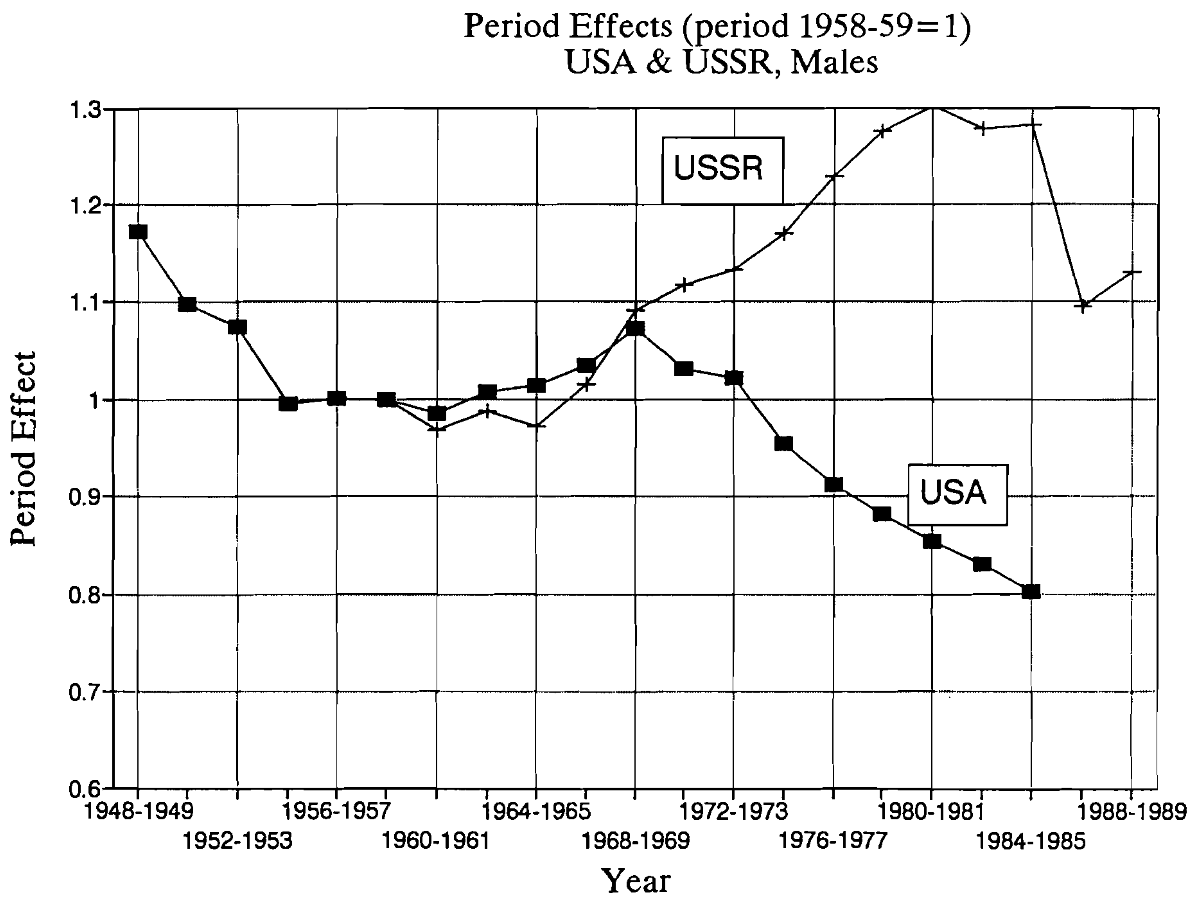
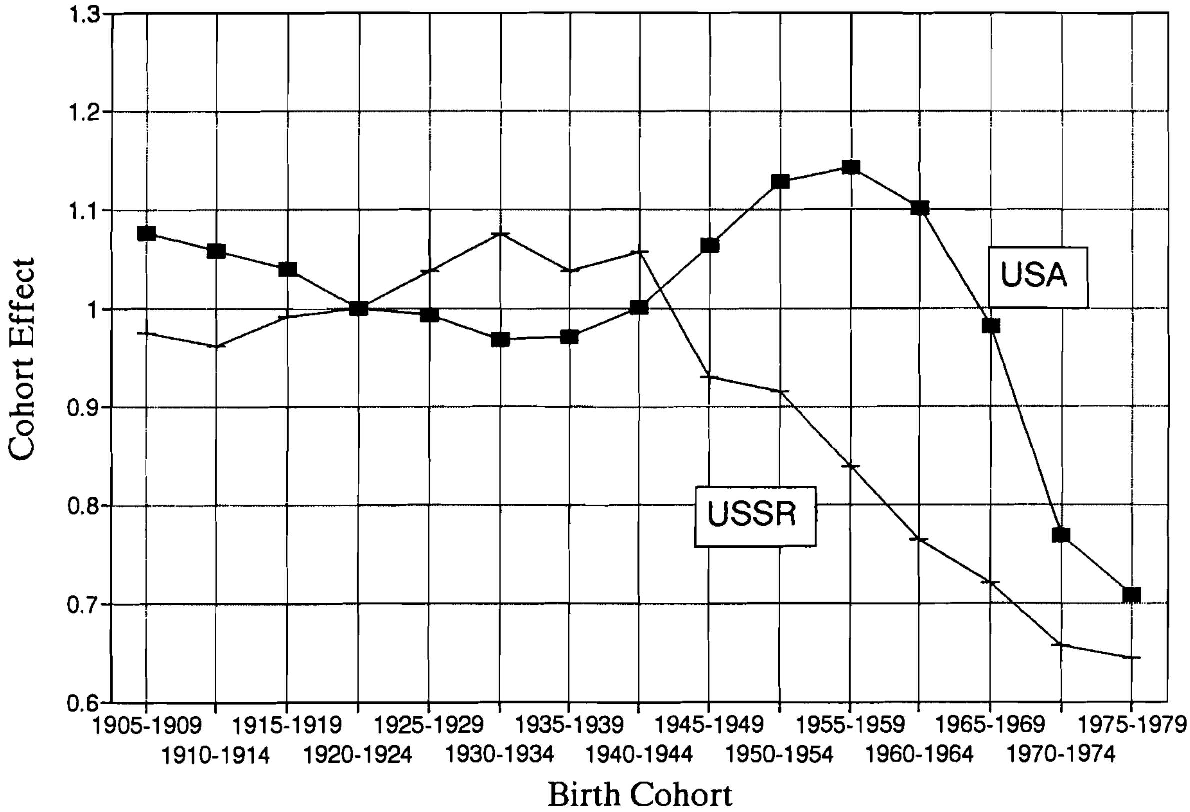
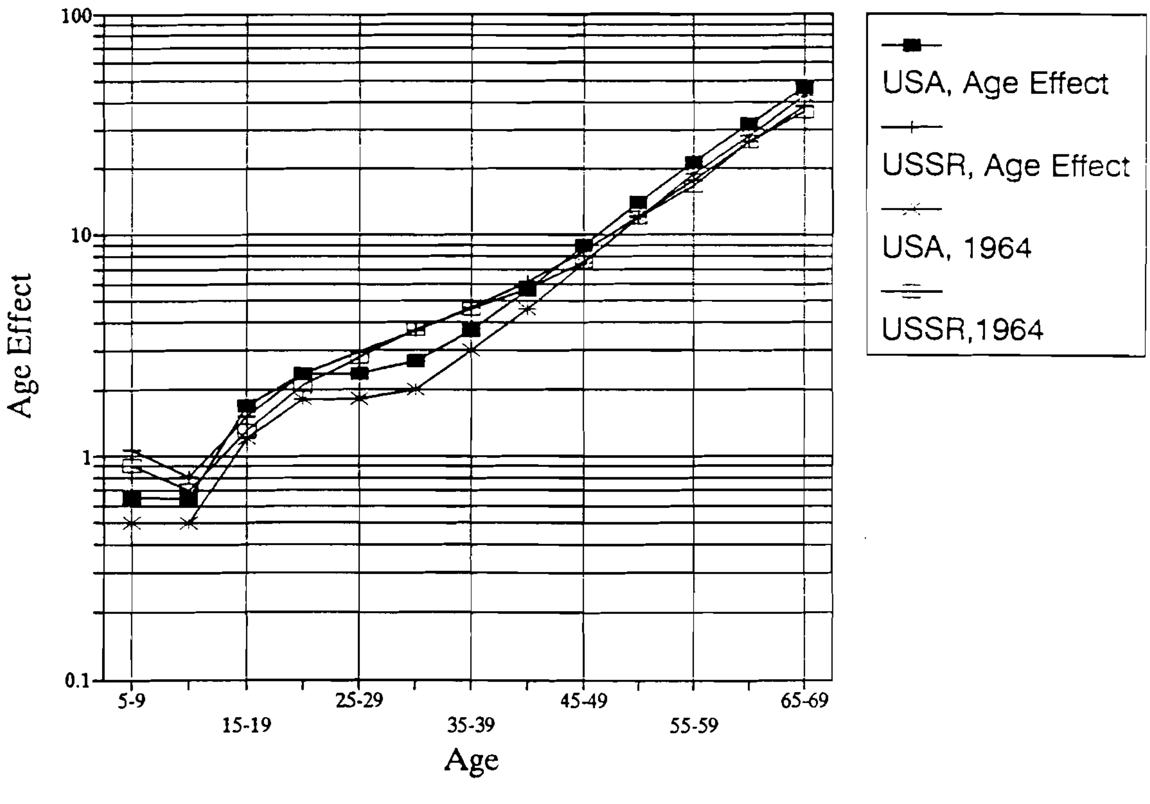


Figure 5a-c. Comparative analysis between USA and USSR.

Cohort Effects (cohort 1920-24=1)
USA & USSR, Males



Age Effects, Ages 5-69 & Data for 1964
USA & USSR, Males



6. CONCLUSION

The objective of this paper was to review research on APC analysis of mortality and to trace the effects of contemporary and historical factors on mortality change in the Soviet Union. Several events occurred in Soviet history that exert a lasting influence on its people. These influences may be captured by an age-period-cohort model, in which the period effects measure the impact of contemporary factors and the cohort effects denote the effects of the past history of individuals that cannot be attributed to age or stage in the life cycle.

Age-period-cohort models are extensively applied in the study of mortality. The statistical properties of the model, and in particular the identification problem that result from a linear dependence between the variables age, period and cohort, received much more attention than alternative specifications to disentangle various factors that give rise to an observed time series of age-specific mortality rates. A number of solutions to the identification problem are reviewed in this paper. It is claimed that the identification problem is not a problem of model specification but a problem of measurement and interpretation.

The statistical theory of the APC model is presented and it is shown that the APC model belongs to the family of generalized linear models. The parameters of the APC model may therefore be estimated using GLIM. They may also be estimated by any package for log-linear analysis that allows for hybrid log-linear models.

The assessment of the effects of age, period and cohort calls for a simultaneous estimation of the three effect parameters. Some authors estimate the effects in stages. They first determine the age and period effects exhibited by the time series of age-specific data. The cohort effects are obtained from the residuals. Anderson and Silver follow this approach in their study of Soviet mortality. The cohort effects that are obtained this way are too small, although the general pattern is revealed. Simultaneous estimation of the age, period and cohort effects indicate that the impact of the war is larger than estimated by Anderson and Silver.

The application of the APC model separates contemporary and historical factors. The main advantage of the model is that it integrates in a single framework the period perspective and the cohort perspective that characterizes many demographic studies. However, it leaves many questions unanswered. We have not been able to determine the size of the interaction between age, period and cohort. Younger cohorts may respond differently to contemporary factors than older cohorts. The aged may be less able to adjust to hardship than young and middle-aged persons. Part of the interactions that are included in the real data, are captured by the cohort or period parameters. Consequently, the main effects of these factors are over- or underestimated. The interpretation of the effects raises other issues that are not fully resolved. The discussion centers on the exact meaning of a cohort effect and, equivalently, the precise measurement of the impact of experiences in the past. Statistical theory will not be able to resolve the issues. Demographic theory, combined with improved model specification and measurement, may indicate the direction to follow in order to improve the ability to assess the contribution of various factors to demographic changes.

BIBLIOGRAPHY

Aitken, M., D. Anderson, B. Francis, and J. Hinde. 1989. *Statistical Modelling Using GLIM.* Clarendon Press, Oxford.

Anderson, B.A. and B.D. Silver. 1989. Patterns of cohort mortality in the Soviet population. *Population and Development Review* 15(3):471-501.

Andreev, E.M. 1990a. Life expectancy and causes of death in the USSR. Paper presented at the International Symposium on Demographic Processes in the USSR in the 20th Century in the Context of the European Experience, Tbilisi, 8-12 October, 1990.

Andreev, E.M. 1990b. Life expectancy and causes of death in the USSR. Pages 90-116 in A.G. Volkov, Ed. *Demographic Processes in the USSR.* Nauka, Moscow (in Russian).

Andvord, K.F. 1921. Is tuberculosis to be regarded from the aethiological standpoint as an acute disease of childhood? *Tubercle* 3:97-116.

Andvord, K.F. 1930. What can we learn by studying tuberculosis by generations? *Norsk Magasin for Laegevidenskapen* 91:642-660.

Attias-Dunfot, C. 1988. *Sociologie des générations.* Presses Universitaires de France, Paris.

Baltes, P.B. and J.R. Nesselroade. 1979. History and rationale of longitudinal research. Pages 1-39 in J.R. Nesselroade and P.B. Baltes, Eds. *Longitudinal Research in the Study of Behavior and Development.* Academic Press, New York.

Baltes, P.B., S.W. Cornelius, and J.R. Nesselroade. 1979. Cohort effects in developmental psychology. Pages 61-87 in J.R. Nesselroade and P.B. Baltes, Eds. *Longitudinal Research in the Study of Behavior and Development.* Academic Press, New York.

Barrett, J.C. 1973. Age, time and cohort factors in mortality from cancer of the vertix. *The Journal of Hygiene* 71:253-259.

Barrett, J.C. 1980. Cohort mortality and prostate cancer. *Journal of Biosocial Sciences* 12:341-344.

Beard, R.E. 1963. Discussion of "Actuarial method of mortality analysis" by B. Benjamin, pp. 38-54. *Proceedings of the Royal Society of London* B, 159:56-64.

Blossfeld, H.P. 1986. Career opportunities in the Federal Republic of Germany: A dynamic approach to the study of life course, cohort, and period effects. *European Sociological Review* 2:208-225.

Boleslawski, L. 1985. Differences in mortality between generations as a result of the world wars. *Studia Demograficzne* 4/82:51071.

Breslow, N.E. 1985. Cohort analysis in epidemiology. Pages 109-143 in A.C. Atkinson and S.E. Fienberg, Eds. *A Celebration of Statistics*. The ISI centenary volume. Springer Verlag, New York.

Breslow, N.E. and N.E. Day. 1985. The standardized mortality ratio. Pages 55-74 in P.K. Sen, Ed. *Biostatistics. Statistics in Biomedical, Public Health and Environmental Sciences.* North Holland, Amsterdam.

Case, R.A.M. 1956. Cohort analysis of mortality as an historical or narrative technique. *British Journal of Preventive and Social Medicine* 10:159-171.

Caselli, G. 1990. The influence of cohort effects on differentials and trends in mortality. Pages 229-249 in J. Vallin, S. D'Souza, and A. Palloni, Eds. *Measurement and Analysis of Mortality. New Approaches.* Clarendon Press, Oxford. (Translation of French edition: *Mesure et analyse de la mortalité. Nouvelles approches.* Presses Universitaires de France, Paris, 1988, pp. 251-272).

Caselli, G. and R. Capocaccia. 1989. Age, period, cohort and early mortality: An analysis of adult mortality in Italy. *Population Studies* 43:133-153.

Caselli, G., J. Vallin, J. Vaupel, and A. Yashin. 1986. L'évolution de la structure par âge de la mortalité en Italie et France depuis 1900: effets de période et effets de génération. *Dossiers et Recherche*, No. 4, INED, Paris.

Caselli, G., J. Duchene, V. Egidi, S. Santini, and G. Wunsch. 1990. A matter of life and death. Methodologies for the life history analysis of adult mortality. Working Paper No. 151. Institut de Demographie, Université Catholique de Louvain, Louvain-la-Neuve.

Derrick, V.P.A. 1927. Observations on (1) errors of age in the population statistics of England and Wales, and (2) the changes in mortality indicated by the national records. *Journal of the Institute of Actuaries* 58:117-146.

Dinkel, R.H. 1985. The seeming paradox of increasing mortality in a highly industrialized nation: The example of the Soviet Union. *Population Studies* 39: 87-97.

Egidi, V., A. Golini, R. Capocaccia, and A. Verdecchia. 1990. Assessing morbidity levels from mortality measurements: The case of cancer in Italy, 1955-1979. Pages 388-405 in J. Vallin, S. D'Souza, and A. Palloni, Eds. *Measurement and Analysis of Mortality. New Approaches.* Clarendon Press, Oxford.

Frome, E.L. 1983. The analysis of rates using Poisson regression models. *Biometrics* 39:665-674.

Frost, W.H. 1939. The age selection of mortality from tuberculosis in successive decades. *American Journal of Hygiene* 30:91-96.

Geddes, A., C. Osmond, A. Barchielli, and E. Buiatti. 1985. Analysis of trends in cancer mortality in Italy 1951-1978: The effects of age, period of birth and period of death. *Tumori* 71:101-110.

Greenberg, B.G., J.J. Wright, and C.G. Sheps. 1950. A technique for analyzing some factors affecting the incidence of syphilis. *Journal of the American Statistical Association* 45:373-399.

Hastings, D.V. and L.G. Berry, Eds. 1979. *Cohort Analysis: A Collection of Interdisciplinary Readings.* Scripps Foundation for Research on Population, Oxford, Ohio, USA.

Heckman, J. and R. Robb. 1985. Using longitudinal data to estimate age, period and cohort effects in earnings equations. Pages 137-150 in W.M. Mason and S.E. Fienberg, Eds. *Cohort Analysis in Social Research.* Springer Verlag, New York.

Hobcraft, J. and W. Gilks. 1984. Age, period and cohort analysis in mortality studies. Pages 245-264 in J. Vallin, J.H. Pollard, and L. Heligman, Eds. *Methodologies for the Collection and Analysis of Mortality Data.* Ordina Editions for IUSSP, Liège.

Hobcraft, J., J. Menken, and S. Preston. 1982. Age, period and cohort effects in demography: A review. *Population Index* 48:4-43. Reprinted in W.M. Mason and S.E. Fienberg, Eds. 1985. *Cohort Analysis in Social Research.* Springer Verlag, New York, pp. 89-135.

Hobcraft, J. et al. 1985.

Hoem, J.M. 1984. Statistical analysis of a multiplicative model and its application to the standardization of vital rates: A review. Stockholm Research Reports in Demography, No. 16, Section of Demography, University of Stockholm.

Holford, T.R. 1980. The analysis of rates and survivorship using log-linear models. *Biometrics* 36:299-306.

Horiuchi, S. 1983. The long term impact of war on mortality: Old age mortality of First World War survivors in the Federal Republic of Germany. *United Nations Population Bulletin* 5.

Kermack, W.O., A.G. McKendrick, and P.L. McKinlay. 1934. Death rates in Great Britain and Sweden: Expressions of specific mortality rates as products of two factors, and some consequences thereof. *Journal of Hygiene* 34:433-457.

Laird, N. and D. Olivier. 1981. Covariance analysis of censored survival data using log-linear analysis techniques. *Journal of the American Statistical Association* 76: 231-240.

Lidell, F.R.K. 1985. Statistical approaches to cohort studies in epidemiology: An overview. Paper presented at the Conference of the International Statistical Institute, Amsterdam.

Mannheim, K. 1928. Das Problem der Generationen. *Kölner Vierteljahresheft für Soziologie* 7(1928-29):154-184 and 309-330.

Manton, K.G. (n.d.) A review and assessment of models for forecasting health status and life expectancy changes in the elderly population.

Manton, K.G., E. Stallard, and J. Vaupel. 1981. Methods for comparing the mortality experience of heterogeneous populations. *Demography* 18.

Mason, W.M. and S.E. Fienberg, Eds. 1985. *Cohort Analysis in Social Research. Beyond the Identification Problem.* Springer Verlag, New York.

Mason, W.M. and H.L. Smith. 1985. Age-period-cohort analysis and the study of deaths from pulmonary tuberculosis. Pages 151-227 in W.M. Mason and S.E. Fienberg, Eds. *Cohort Analysis in Social Research.* Springer Verlag, New York.

Mayer, K.U. and J. Huinink. 1990. Age, period, and cohort in the study of the life course: A comparison of classical A-P-C analysis with event history analysis or farewell to Lexis? Pages 211-232 in D. Magnusson and L.R. Bergman, Eds. *Data Quality in Longitudinal Research.* Cambridge University Press, Cambridge.

McCullagh, P. and J.A. Nelder. 1983. *Generalized Linear Models.* Chapman and Hall, London (2nd edition, 1989).

McKeown, T. 1976. *The Modern Rise of Population.* Academic Press, New York.

Mortana, G. 1912. Le variazioni della mortalitàde generazione a generazione in Svezia. *Annali di Statistica*, Serie V, 4.

Nelder, J.A. 1984. Models of rates with Poisson errors. *Biometrics* 40:1159-1160.

Nelder, J.A. and R.W.M. Wedderburn. 1972. Generalized linear models. *Journal of the Royal Statistical Society* A, 135:370-384.

Osmond, C. and M.J. Gardner. 1982. Age, period and cohort models applied to cancer mortality rates. *Statistics in Medicine* 1:245-259. (+ history of APC analysis of mortality rates)

Osmond, C. and M.J. Gardner. 1989. Non-overlapping cohorts don't resolve the identification problem. *American Journal of Epidemiology* 129(1):31-35.

Osmond, C., M.J. Gardner, and E.D. Acheson. 1982. Analysis of trends in cancer mortality in England and Wales during 1951-1980 separating changes associated with period of birth and period of death. *British Medical Journal* 284 (3 April):1005-1008.

Pollard, J.H. 1987. Projection of age-specific mortality rates. *Population Bulletin of the United Nations* 21/22:55-69.

Preston, S. and E. van de Walle. 1978. Urban French mortality in the Nineteenth Century. *Population Studies* 32.

Robertson, C. and P. Boyle. 1986. Age, period and cohort models: The use of individual records. *Statistics in Medicine* 5:527-538.

Ryder, N.B. 1965. The cohort in the study of social change. *American Sociological Review* 30:843-861. Reprinted in W.M. Mason and S.E. Fienberg, Eds. 1985. *Cohort Analysis in Social Research.* Springer Verlag, New York, pp. 9-44.

Tu, E.J.C. and J.L.C. Chuang. 1983. Age, period and cohort effects on maternal mortality: A linear logit model. *Social Biology* 30(4):400-412.

Vallin, J. 1973. La mortalité par génération en France depuis 1899. *Travaux et Documents, Cahier 63.* INED, Paris.

Vallin, J. 1984. Tables de mortalité du moment et par génération 1899-1981: mise à jour provisoire de tables annexées du *cahier 63.* INED, Paris.

van Poppel, F.W.A. 1990. Determinants and consequences of mortality in the Netherlands: Some suggestions for research. Pages 233-275 in C.A. Hazeu and G.A.B. Frinking, Eds. *Emerging Issues in Demographic Research.* Elsevier Science Publishers, Amsterdam.

Willekens, F.J. and N. Baydar. 1986. Age-period-cohort models for forecasting fertility. Working Paper No. 45. Netherlands Interdisciplinary Demographic Institute, The Hague.

Wilmoth, J., J. Vallin, and G. Caselli. 1988. When does a cohort's mortality differ from what we might expect? Paper presented at the Annual Meeting of the Population Association of America, New Orleans, Louisiana, USA.

APPENDIX A

The measurement issue in APC analysis, illustrated with the Lexis diagram

The objective of APC analysis of a time series of age profiles (age-specific data) is to isolate the effects of the life course, contemporary factors and historical factors. The stages in the life course are approximated by age, the contemporary factors by period, and the historical factors by cohort (year or period of birth or another event-origin).

The separation of age, period and cohort effects requires proper measurement of the age at which the event occurs, the year (period) in which it occurs and the cohort to which the person experiencing the event belongs. Most data that are available for APC analysis lack one of the three variables that characterize the timing of the event. In general, the data are classified by age and period (age-period tables). Cohort experiences are inferred from the diagonals of the table. This approach is the basis of the identification problem. Because the year of birth is unknown for an individual of a given age who experiences an event in a given year, the cohort effect cannot be identified unambiguously.

The Lexis diagram is a two-dimensional diagram locating the events with reference to

- the date of occurrence of the event of interest,
- the date of occurrence of the event-origin (e.g. birth), and
- the duration since the event-origin (e.g. age).

If each individual is under continuous observation, the timing of events can be measured precisely and for each individual a lifeline can be drawn. Figure A1 shows four lifelines a, b, c and d. Consider lifeline c and let P denote an event occurring at exact age t to a person of exact age x. The individual to which the lifeline c refers is born at exact time t-x. Note that when any two of the three time measures are known, the third can be determined precisely. The age, period and cohort variables are therefore linearly dependent:

$$c = t - x$$

where x denotes age, t period and c cohort.

Even if all events are recorded by a continuous-time observation, the data frequently available to demographers are grouped data. The grouping may be over time, cohort and/or age, yielding discrete time, cohort and/or age intervals. The grouping generally results in intervals of one or five years. A consequence of time grouping is that the location of the event on the lifeline is only known approximately and that we cannot infer any more the exact value of a time variable from knowledge of the two other variables. We may restore the relation if we measure the age, period and cohort intervals.

The discrete time or observation intervals can be visualized in the Lexis diagram. The figure shows an age interval (x,x+1), a period interval (t,t+1) and a cohort interval (t-x-1,t-x), The cohort consists of the group of people who experienced the event-origin in the time period from t-x-1 to t-x. Note that the cohort may be identified by either the year of event-origin (birth) or the age in completed years at time t or t+1. The timing of the event is generally measured by two of the three time variables. The further analysis depends on how the timing is measured. Four cases, known as observation plans, may be distinguished:

A. Cohort (cohort-age) observation:

A cohort observational plan records for a person experiencing an event the cohort to which the person belongs and the seniority in completed years at the time of the event (parallelogram WQSP). The observation interval extends over two calendar years.

B. Period-cohort observation:

A period-cohort observational plan records the calendar year in which the event occurs as well as the cohort to which the person belongs (parallelogram PQRS). The observational plan extends over two cohorts.

C. Period (period-age) observation:

A period observational plan records the calendar year in which an event occurs as well as the seniority of the person in completed years at the time of the event (square PQSV). The period observation interval covers two cohorts.

D. Age-period-cohort observation:

An age-period-cohort observation plan records the calendar year in which the event occurs as well as the seniority of the person in completed years at the time of the event and the cohort to which the person belongs (triangle PQS). The APC observation interval extends over only one period, one age and one cohort. Data presented by age, period and cohort are frequently referred to as doubly classified data.

Most often a demographic time series is represented as a time series of period-age observations. Regarding this representation two remarks can be made:

1. The events (or rates) pertaining to a given age and period category cover the experience of two cohorts.
2. The diagonal sequence of age-by-period classified data fail to cover all the experience of any of these two cohorts.

The first remark can easily be detected in the Lexis diagram. The age-by-period scheme given in the square PQSV covers the events occurring to cohorts t-x-1 and t-x. Hence the relationship "cohort = period - age", that is assumed to be inherent to age, period and cohort as classification variables, does not hold. Strictly speaking, the cohort cannot be predicted from age-period data. The second remark points to the fallacy of the assumption, implicit in most APC analysis, that the diagonal sequences of age-period data suffice to study cohort experiences. Two-factor classified data fail to provide accurate information on the third factor. That is the reason why the identification problem is a measurement problem rather that a model specification problem that is inherent to APC models.

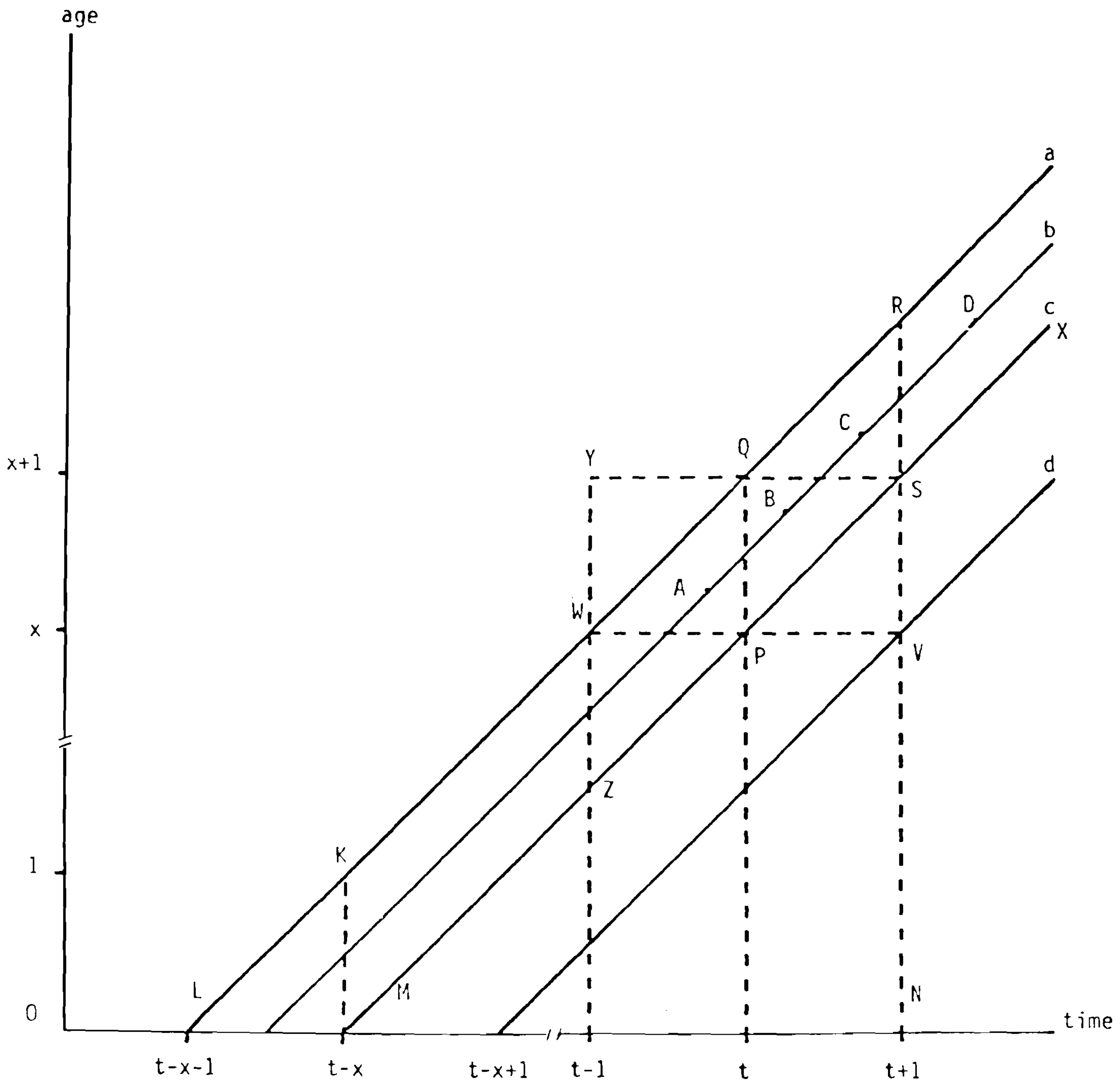


Figure A1. Lexis diagram.

APPENDIX B

Additional references on APC analysis of mortality trends

Boyle, P. and C. Robertson. 1987. Statistical modelling of lung cancer and laryngeal cancer incidence in Scotland, 1960-1979. *American Journal of Epidemiology* 125:731-744.

Letters to the Editor:
Boyle, P. and C. Robinson. 1989. *AJE* 129(1):225-226.
Tango, T. 1988. *AJE* 128:677-678.

Boyle, P., N.E. Day, and K. Magnus. 1983. Mathematical modelling of malignant melanoma trends in Norway, 1953-1978. *American Journal of Epidemiology* 118:887-896.

Breslow, N.E., J.H. Lubin, P. Marek, and B. Langholz. 1983. Multiplicative models and cohort analysis. *Journal of the American Statistical Association* 78:1-12.

Clayton, D. and E. Schifflers. 1987. Models for temporal variation in cancer rates: Age-period, age-cohort models. *Statistics in Medicine* 6:449-467.

Collins, J.J. 1982. The contribution of medical measures to the decline of mortality from respiratory tuberculosis: An age-period-cohort model. *Demography* 19:409-427.

Day, N.E. and B. Charnay. 1982. Time trends, cohort effects and ageing as influence on cancer incidence. Pages 51-62 in K. Magnus, Ed. *Time Trends in Cancer Incidence: Causes and Practical Implications.* Hemisphere Publishing Company, New York.

Decarli, A. and C. La Vecchia. 1987. Age, period and cohort models: Review of knowledge and implementation in GLIM. *Revista di Statistica Applicata* 20:397-410.

Glenn, N.D. 1976. Cohort analysts' futile quest: Statistical attempts to separate age, period and cohort effects. *American Sociological Review* 41:900-904.

Knoke, D. and M. Hout. 1976. Reply to Glenn. *American Sociological Review* 41:905-908.

Mason, W.M., K.O. Mason, and H.H. Winsborough. 1976. Reply to Glenn. *American Sociological Review* 41:904-905.

Goldstein, H. 1979. Age, period and cohort effects--a confounded confusion. *Bulletin in Applied Statistics* 6:19-24.

Hagenaars, J.A. and N.P. Cobben. 1978. Age, cohort and period: A general model for the analysis of social change. *The Netherlands Journal of Sociology* 14:59-91.

Hoff, N. M. van Het. 1979. Cohort analysis of lung cancer in the Netherlands. *Journal of Epidemiology* 8(1):41-47.

Hogan, D.P. 1984. Cohort comparison in the timing of life events. *Developmental Review* 4:289-310.

Holford, T.R. 1983. The estimation of age, period and cohort effects of vital rates. *Biometrics* 39:311-324.

Holford, T.R. 1985. An alternative approach to statistical age-period-cohort analysis. *Journal of Chronic Disease* 38(10):831-836.

Holman, C.D., I.R. James, P.H. Gattey, and B.K. Armstrong. 1980. An analysis of trends in mortality from malignant melanoma of the skin in Australia. *International Journal of Cancer* 26:703-709.

Holman, C.D., I.R. James, M.R. Segal, and B.K. Armstrong. 1981. Recent trends in mortality from prostate cancer in male populations of Australia and England and Wales. *British Journal of Cancer* 44:340-348.

Huinink, J. 1988. Methoden der explorativen Kohortenanalyse. *Zeitschrift für Bevölkerungswissenschaft* 14(1):69-87.

James, I.R. and M.R. Segal. 1982. On a method of mortality analysis incorporating age-year interaction. *Biometrics* 38:433-443.

Kupper, L.L., J.M. Janis, and A. Karmous, et al. 1985. Statistical age-period-cohort analysis: A review and critique. *Journal of Chronic Disease* 38:811-840 (including discussion).

Mayer, K.U. and J. Huinink. 1987. Age, period and cohort in the study of the life course: A comparison of classical A-P-C analysis with event history analysis or farewell to Lexis? Paper presented at the Workshop on Methodological Issues in Longitudinal Research, Rönneberga, Stockholm, December 13-16.

Selvin, S. and S. Sachs. 1979. Analysis of cohort effects from cross-sectional data. *Comp. Progr. Biomed.* 9:218-222.

Tu, E. J-C. 1982. *Age, Period and Cohort Effects on Maternal Mortality.* Chuang Jane-Lan, Washington, D.C.