MPIDR-NES Training Programme

Moscow, New Economic School,14th January - 1st February 2013

Population and Health

Лекция 5. Изменения смертности в зависимости от времени, возраста и когорты

Lecture 5. Mortality across time, age, and cohort

MAX PLANCK INSTITUTE FOR DEMOGRAPHIC RESEARCH

MAX-PLANCK-INSTITUT FÜR DEMOGRAFISCHE FORSCHUNG

РЭШ
Российская экономическая школа

Outline of the Lecture

- General considerations
- Mortality curve and historical health progress
- Age-specific mortality change with time and by country and its presentation by the Coale-Demeny MLTs
- Mortality surfaces
- Modeling of mortality as a function of age
- Cohort effects in mortality - echoes of the past
- False cohort effects

General considerations - I

Mortality hazard can be considered as a function of numerous variables. Among them – demographic, biological, ethno-cultural, socio-economic, medical, and many other categories.

Three principal characteristics considered in classic demography – age, calendar time, and year of birth (cohort).

Age is a fundamental characteristic of human beings that also determines biologically significant periods of human life: infancy, childhood, adolescence, maturity/reproductive life, and aging. Each of these periods is characterized by typical health risks (e.g. congenital abnormalities in infancy, infections in childhood, health problems related to pregnancy and childbirth and accidents during active reproductive and working life, diseases related to age-dependent degeneration of the body in older age).

General considerations - II

Mortality changes **by calendar year** tell about changeable health conditions in different time periods. In most of cases, mortality hazard becomes lower with time. Rates of mortality improvement or deterioration are crucial for evaluation of the public health and social situations in a country.

Year of birth determines generations of humans (e.g. **cohorts** of individuals born in the same year) and their joint destinies over the life course. For example, males born in Russia and Germany in the early 1920s were exposed to a high risk of premature death in World War 2. Generations that experience famine in early life can be exposed to higher risk of cardiovascular disease and/or cancer in later life.

Variables *Age*, *Year*, and *Cohort* are inter-dependent:

Year = Cohort + Age

Mortality curve: a general decrease and stability of the shape

Age-specific death rates (single-year ages) in Sweden in 1850, 1900, 1930, 1950, 1970, and 2000

In spite of the overall mortality lowering, mortality curve keeps its principal features:

- relatively high mortality at age 0 and extremely rapid mortality fall at ages between 0 and 11-13 years;

- some elevation at young adult ages due to accidents, violence, and maternal causes among women;

- exponential mortality increase over a wide range of ages from 30 to 80 or 85;

- a slow-down in mortality increase at ages 90+.

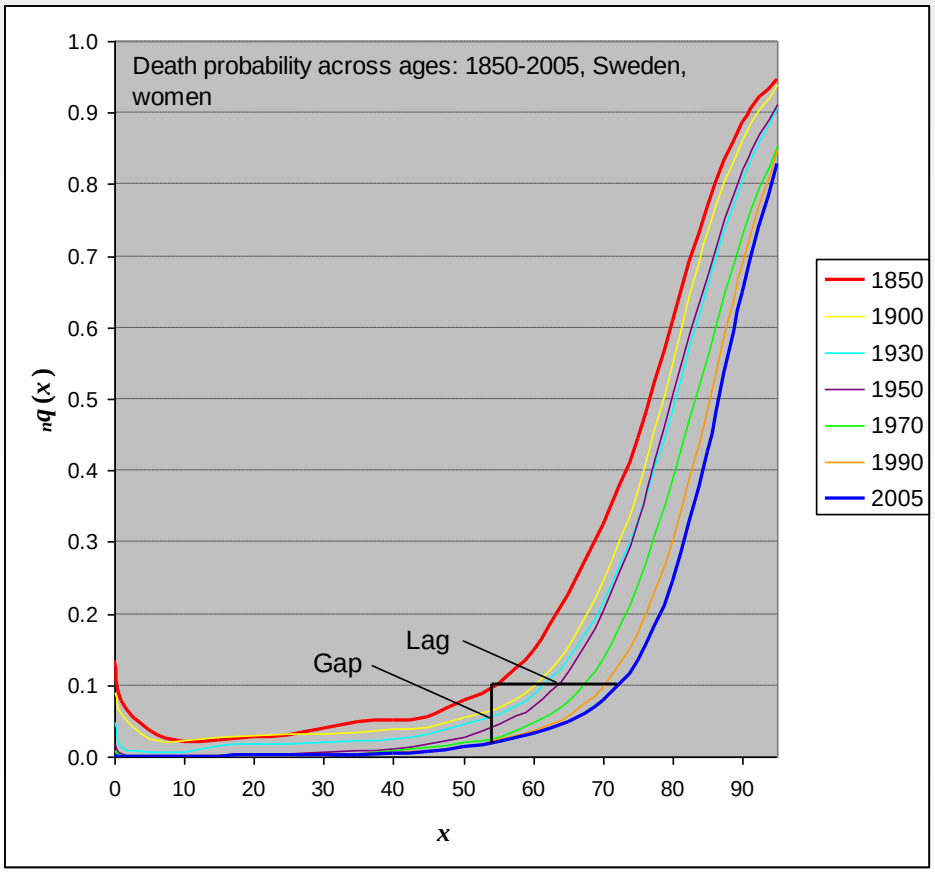
Why such a shape of the mortality curve and why it is always the same ?

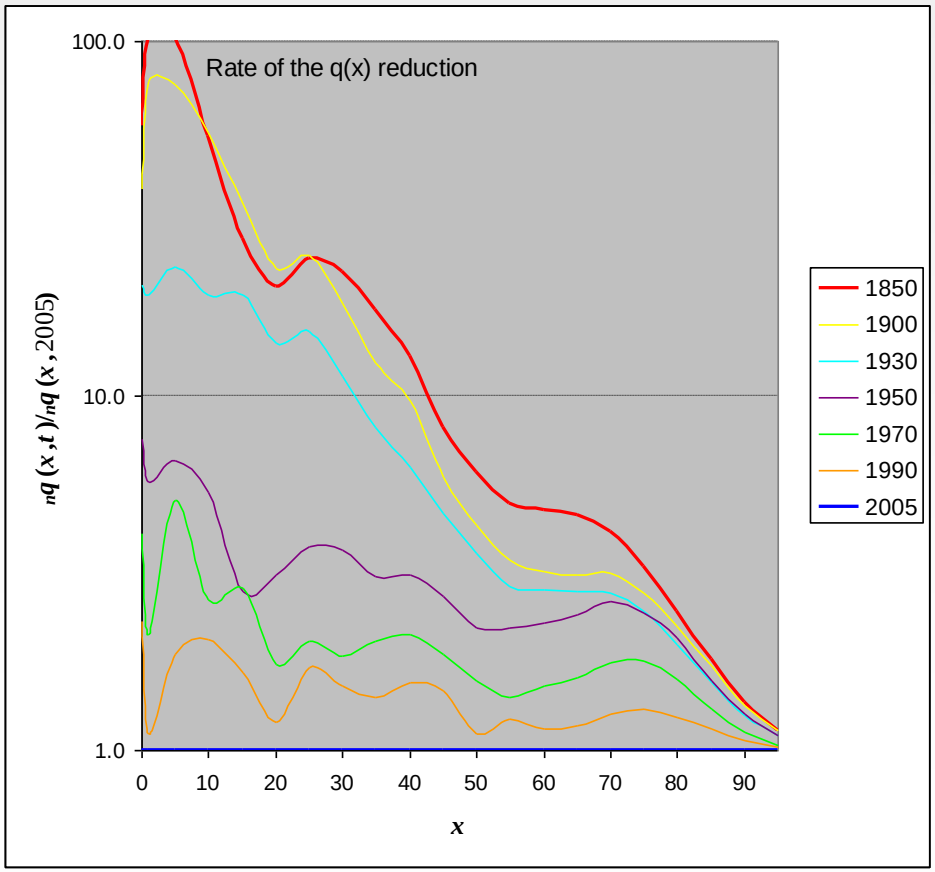
Mortality curve as a whole lowers with time due to the historical health progress. However, its shape sustains.

Most likely, the shape of the mortality curve reflects certain fundamentals of the human biology.

- High mortality just after birth reflects a “potential barrier” on the way to further life. Babies with low birth weight, congenital abnormalities, poor nutrition and contacts to infections have lower chance for survival.
- Later in childhood, mortality decreases and reaches a minimum since the childhood risks are becoming lower while the adult risks are not there yet.
- At transitional and young adult ages from 15 to 40, people are mostly exposed to risks related to childbearing (maternal death among women) as well as to risky behaviors (more among men). At these ages, many deaths are caused by suicides, various accidents, and alcohol drinking.
- Exponential mortality increase over ages from 40 to 80 reflects the process of aging. In the 1960s, Strehler and Mildvan showed that a linear deterioration in functioning of the organism’s sub-systems leads to a population’s mortality hazard that increases exponentially with age.

A closer look at age-specific mortality changes

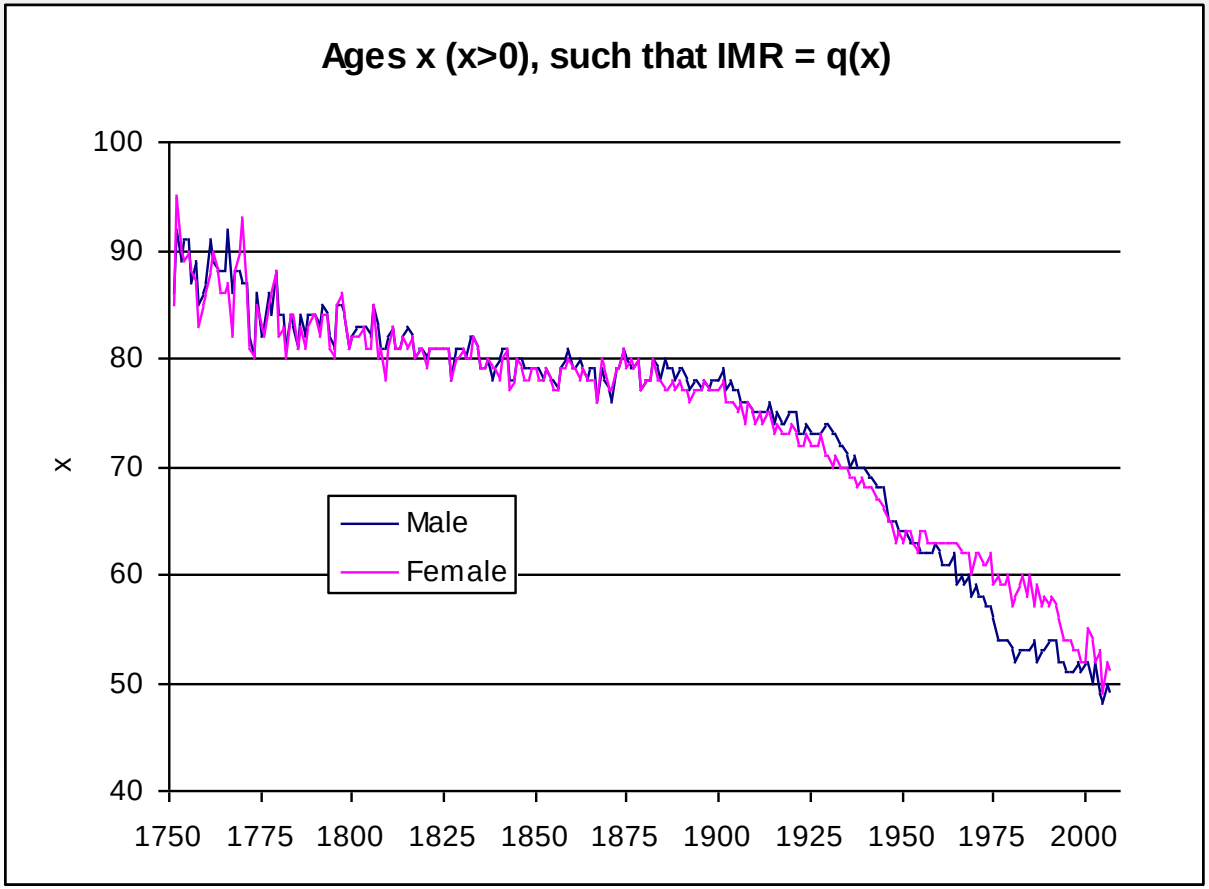




Age-specific death probabilities and their proportional changes: Swedish females, 1850-2005.

Apart from the general mortality decrease, mortality reductions are uneven with respect to age. Greater reductions at young ages. Growing importance of midlife and old-age mortality after the 1930s.

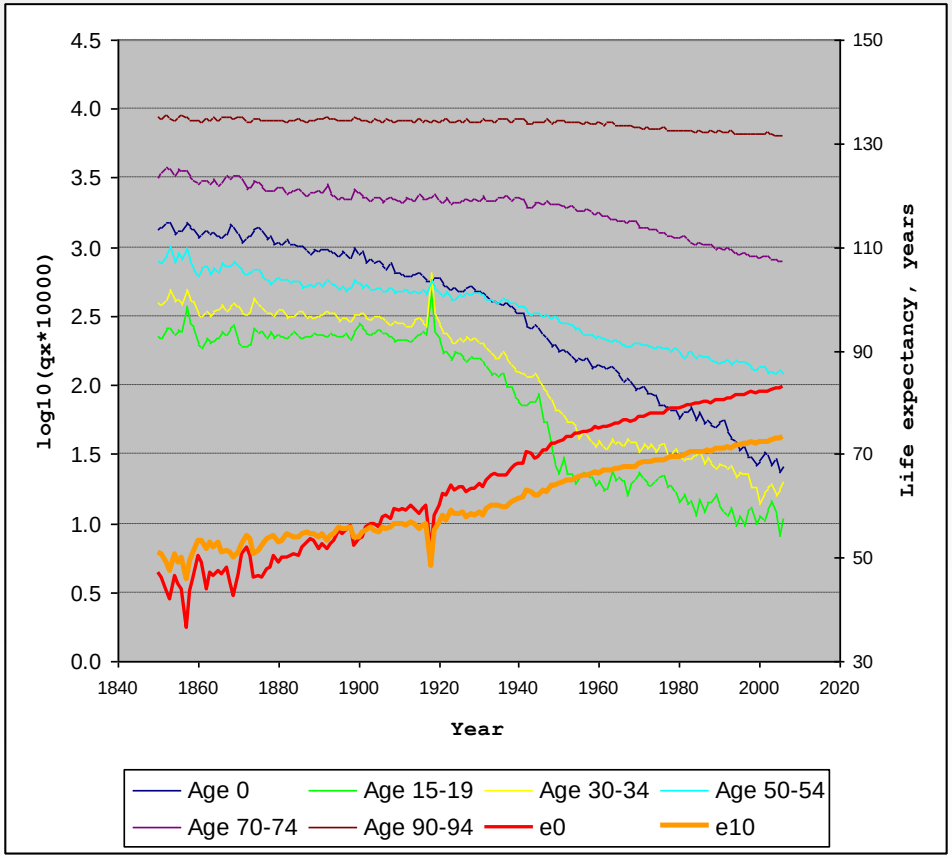
U-shape curve transforms into a J-shape curve
Swedish males and females, since 1750



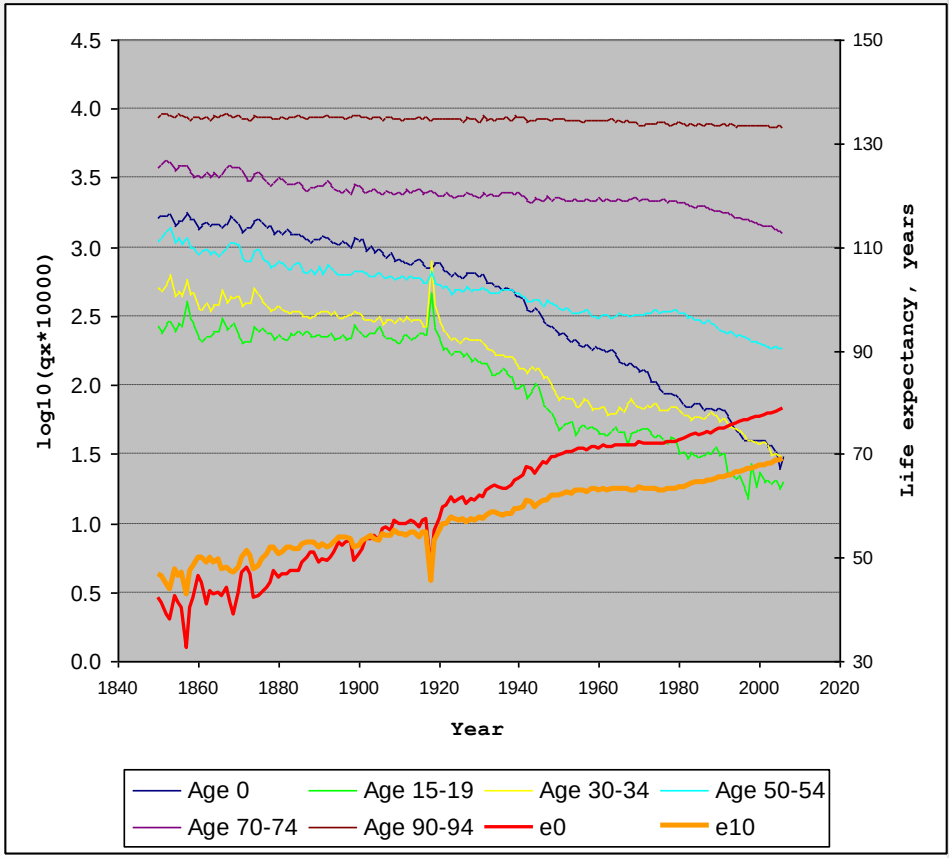
Particularly steep mortality decrease in infancy and childhood leads to a lowering of the left branch of the mortality curve compared to its right branch (U →J).
Around 1750, IMR is equal to q(x) at age 90. In 2005, IMR corresponds to q(x) at age 50.

Differential but uniform changes in q_x, e_0, and e_{10}

Females



Males



Trends in age-specific death probabilities and in life expectancies at ages 0 and 10: Sweden, 1850-2005.

Uniformity of the trends – there is always a decrease that mirrors increases in e_0 and e_{10}

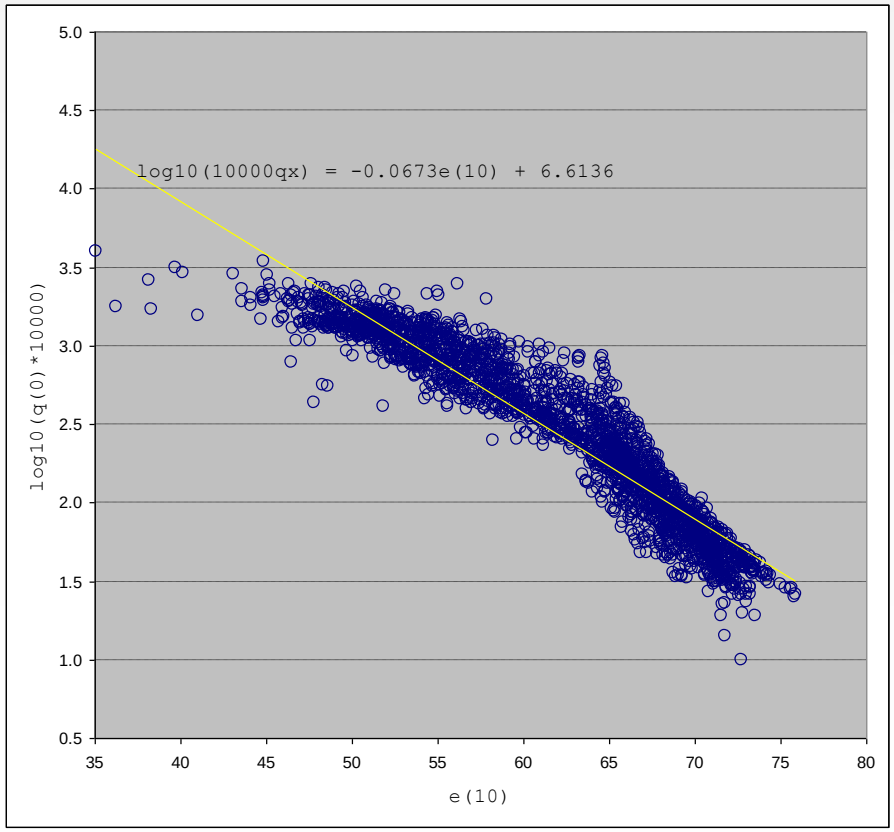
Inter-correlations among q_x, e_0, and e_{10} series for females in Sweden, 1850-2005

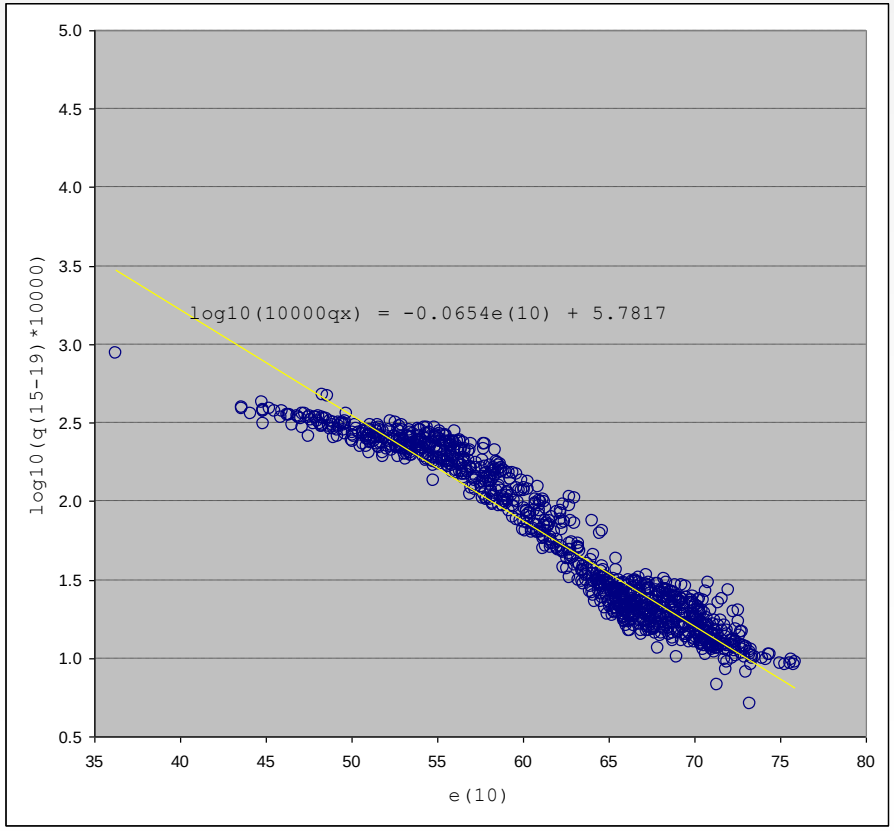
Females	1q0	5q15	5q25	5q35	5q45	5q55	5q65	5q75	5q85	5q90
1q0	1									
5q15	0.8866	1								
5q25	0.8937	0.9874	1							
5q35	0.9669	0.9598	0.9678	1						
5q45	0.968	0.9232	0.9296	0.9832	1					
5q55	0.9603	0.8697	0.8827	0.9566	0.9873	1				
5q65	0.9397	0.8382	0.8545	0.9312	0.9694	0.99	1			
5q75	0.8790	0.8038	0.8146	0.8754	0.9204	0.9446	0.9734	1		
5q85	0.7846	0.7613	0.7673	0.8004	0.8433	0.8617	0.9065	0.9725	1	
5q90	0.7556	0.7533	0.7552	0.7785	0.8180	0.8336	0.8791	0.9502	0.9870	1
	q0	q15	q25	q35	q45	q55	q65	q75	q85	q90
e0	-0.9807	-0.9372	-0.9410	-0.9835	-0.9828	-0.9685	-0.9565	-0.9233	-0.8573	-0.8351
e10	-0.9403	-0.9475	-0.9521	-0.9732	-0.9754	-0.9604	-0.9573	-0.9466	-0.9060	-0.8887

Temporal changes are highly correlated among ages. Even probabilities of dying at extreme ages 0 and 90-94 are strongly correlated (Pearson's r=0.7556). At all ages, changes in probabilities of dying are highly and negatively correlated with e_0 and e_{10}.

Variation across time (since 1850) and country (25 countries)

Linear relationships between log-probabilities of death at different ages and e_{10} : 2657 female life tables from the HMD. (# 1: ages 0 and 15-19).

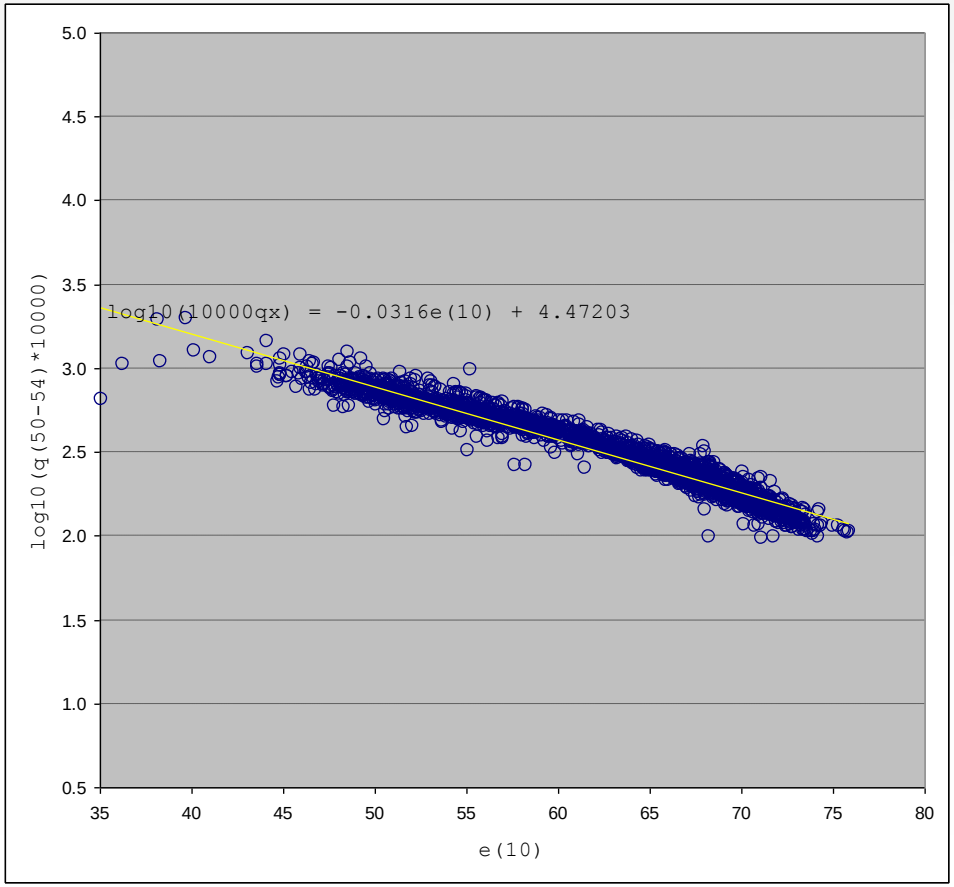


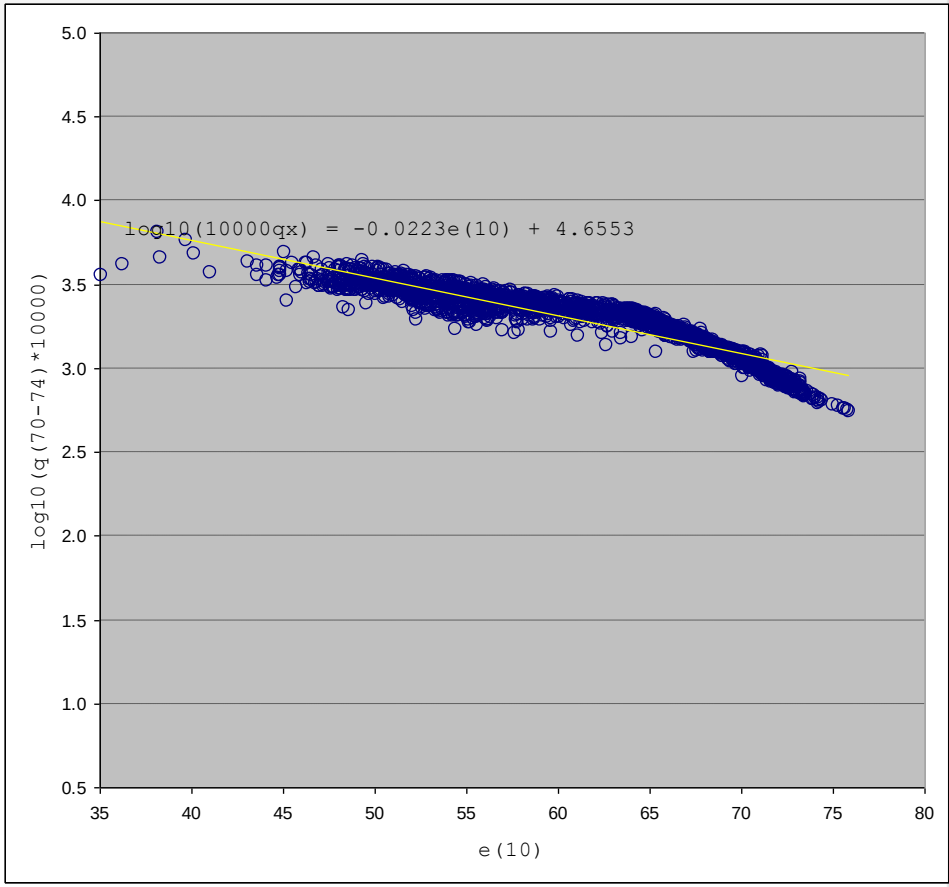


The set of 25 countries includes all HMD countries except Bulgaria, Hungary, and countries of the former Soviet Union.

Variation across time (since 1850) and country (25 countries)

Linear relationships between log-probabilities of death at different ages and e_{10} : 2657 female life tables from the HMD. (# 2: ages 50-54 and 70-74).





The empirical regularities

The scatter plots on the previous two slides show that at all ages female age-specific mortality is strongly and nearly linearly associated with e_{10}. The same can be shown for males, but respective relationships are slightly weaker.

Thus, the variation across time and space has the same and (relatively) simple character that is comparable to the time trends within one country (Sweden). Although, slopes of the OLS regressions are greater at younger than at older ages, at every age there is a strong relationship between the probability of dying and e_{10}.

This way we come to the central idea of the Coale-Demeny Model Life Tables (MLTs): at every age, age-specific probability of dying can be estimated from only one value – life expectancy at age 10.

How the Coale-Demeny MLTs are built

A simplified description of the C-D MLTs of 1984.

1.Collect a large number of accurate human life tables (abridged LTs) – 326 LTs.

2.Pre-classify these life tables into regional families according to (somewhat) different shapes of the mortality curves (West, East, North, South).

3.Within each family, for age groups x estimate regression coefficients A_x and B_x, such that ${}_nq_x = A_x + B_x e_{10}$

4.Build model female life tables from series of e_{10} values beginning from e_{10}=20 years to e_{10}=80 years.

5.Estimate male e_{10} values corresponding to respective female values using certain simple relationship between male and female life expectancies.

6.Estimate male model life tables within each regional family (the same way as in step 3) using the male coefficients A_x and B_x and the male e_{10} series.

MLTs are used as typical (“model”) age patterns of mortality and for estimation of age-specific mortality from mortality at a few ages in countries with no mortality statistics.

Mortality surface – a graphic image of mortality

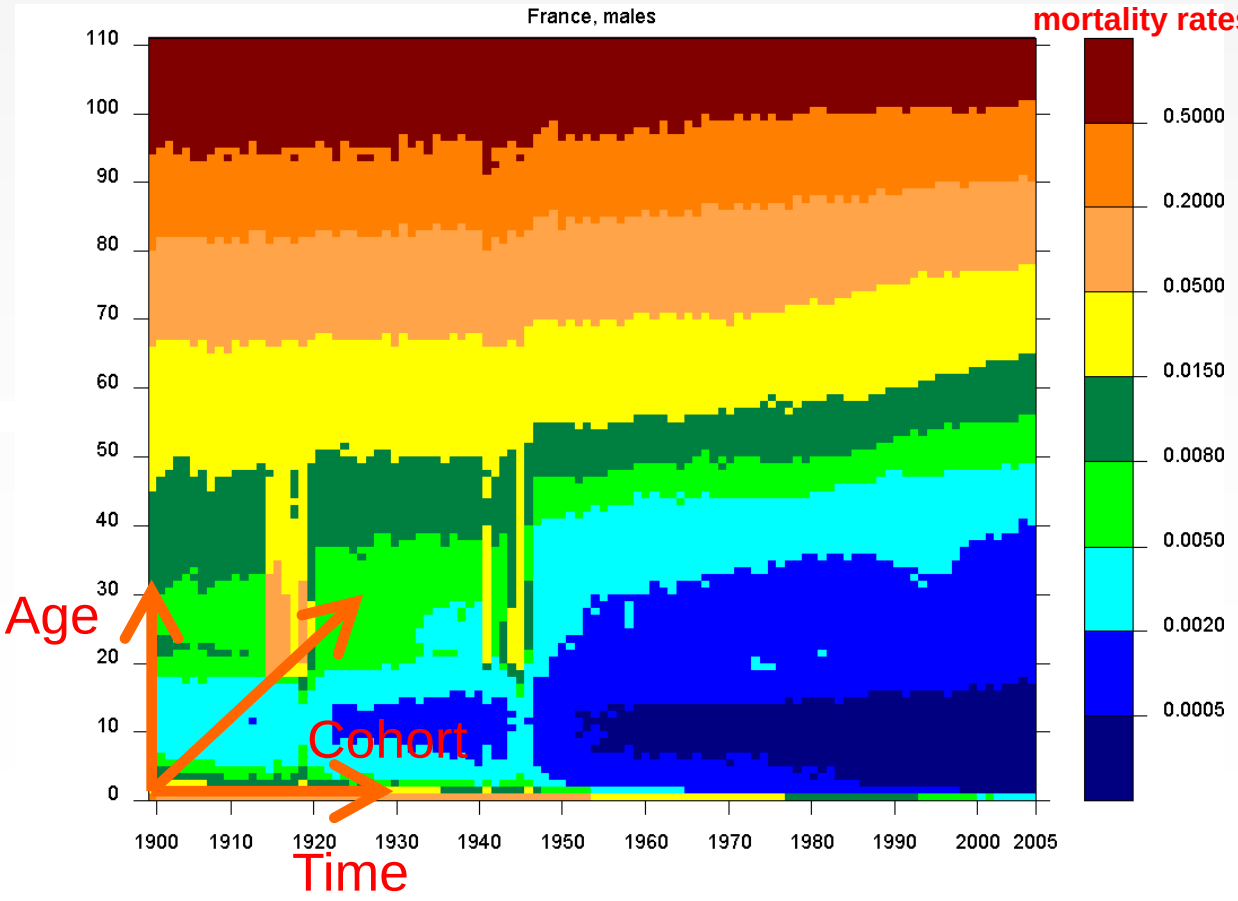
As you know, mortality rate is strongly depends on age and time. Birth cohort is an additional dimension that can be expressed as a function of age and time:

Year_of_birth=Time – Age

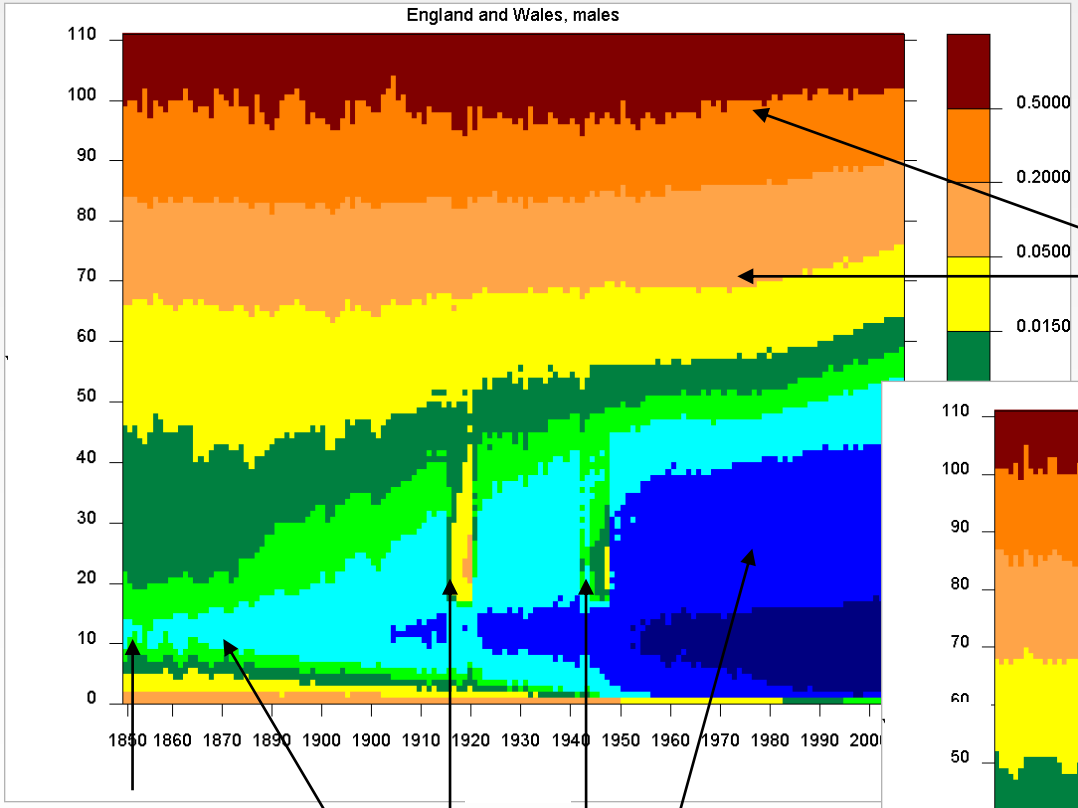
Mortality surface is an instrument for visual inspection of mortality variation over Age, Time, and Cohort.

One can see the same features:

- Mortality is relatively high at the very beginning of life.
- Mortality strongly increases with age after age 30-40.
- Mortality generally decreases with time.
- Sometimes mortality fluctuates due to wars and epidemics.



A few more observations: mortality surface for England and Wales since 1850

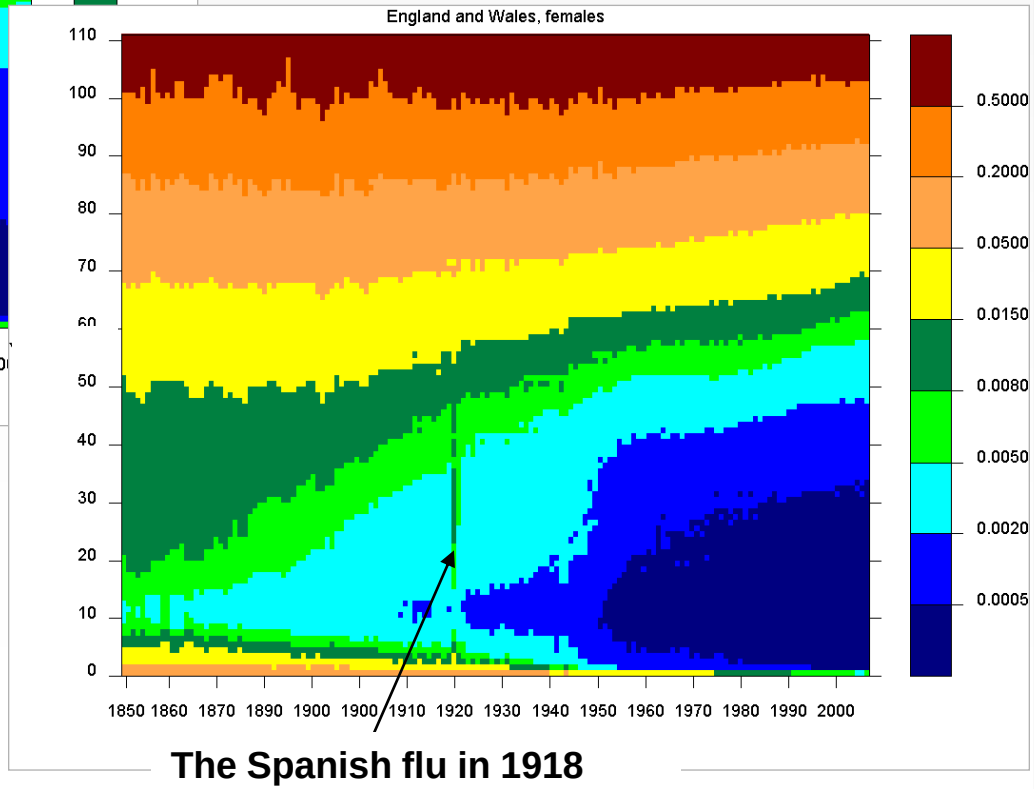


Progress in old-age mortality since the 1970s

Mortality fluctuations due to epidemics

WW1 and WW2

Continuous progress at young ages



The Spanish flu in 1918

Parametric description of mortality age patterns. The Gompertz model

Parametric models of mortality curves are useful for building these curves from fragmentary data in countries with poor data availability (or quality) and also for studying age patterns of mortality.

$$\mu(x) = R \cdot e^{\alpha x}$$

R - a baseline level of mortality,
α – steepness of mortality increase with age

On the discrete LT data: the model is as follows

$$M_x = R \cdot e^{\alpha \bar{x}}$$

$\bar{x}$ – is the mean age at death within elementary age interval [x, x+n)

The two parameters can be estimated from a linear OLS regression:

$$\ln(M_x) = \ln R + \alpha \cdot \bar{x}$$

Note: Gompertz model can be applied to the Mx or qx curves at ages above 25-35 (up to 85-90).

AGE. The Gompertz-Makeham model

The Gompertz-Makeham model includes an additional parameter *A* responsible for „exogenous“ mortality from acute causes that can kill people regardless their age.

$$M_x = A + R \cdot e^{\alpha \bar{x}}$$

The three parameters can be estimated in the following way:

Step 1: estimate the two Gompertz parameters from the same OLS regression;

$$\ln(M_x) = \ln(R) + \alpha \cdot \bar{x}$$

Step2: estimate the Makeham parameter A as an average of age-specific residuals of this regression

$$A = \frac{1}{N} \sum_x (M_x - R \cdot e^{\alpha \cdot \bar{x}})$$

N – is the number of age groups

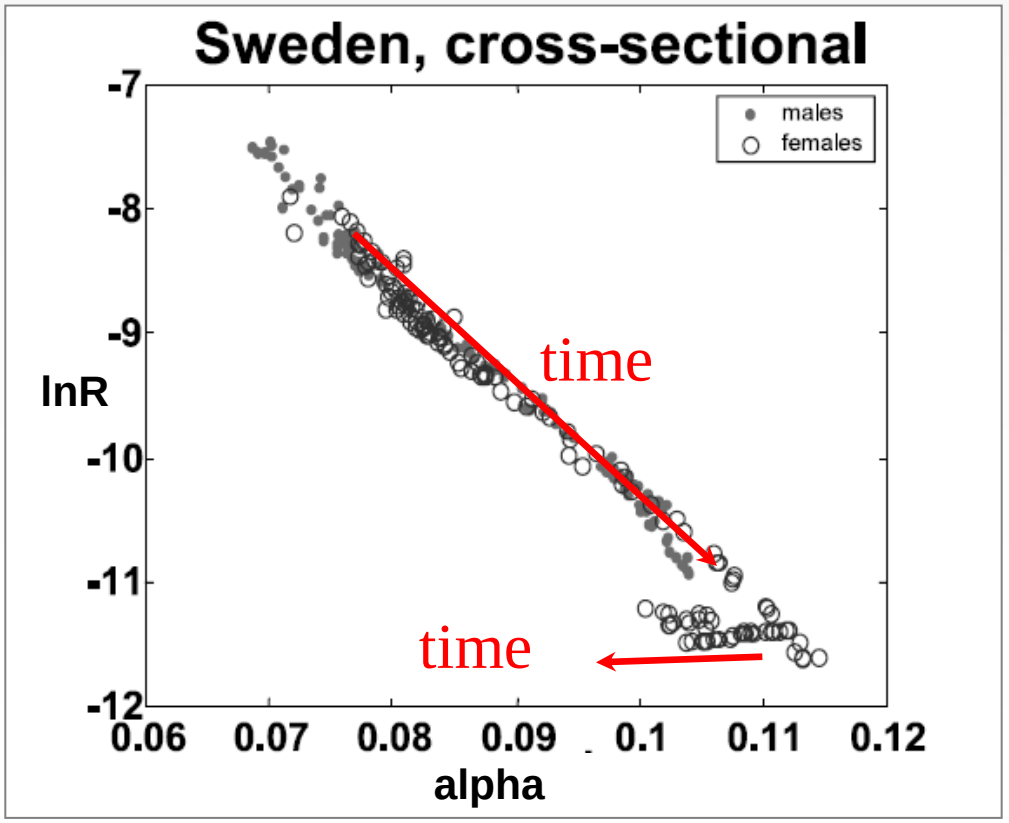
Calculation procedure and examples are given in the following Excel file

Gompertz.xls

Historical evolution of parameters *R* and *Alpha*

The Gompertz model has been connected to certain theories of biological aging. The Strehler-Mildwan's theory (1960) connects values of parameters *R* and *Alpha* with quicker or slower decline in the organism's biological capacity with age.

The S-M framework predicts a negative correlation between ln(*R)* and *Alpha*. This correlation can be observed during very long time, but it has begun to change since the 1950s (Yashin et al., 2002).



Modeling the whole mortality curve: Heligman-Pollard model

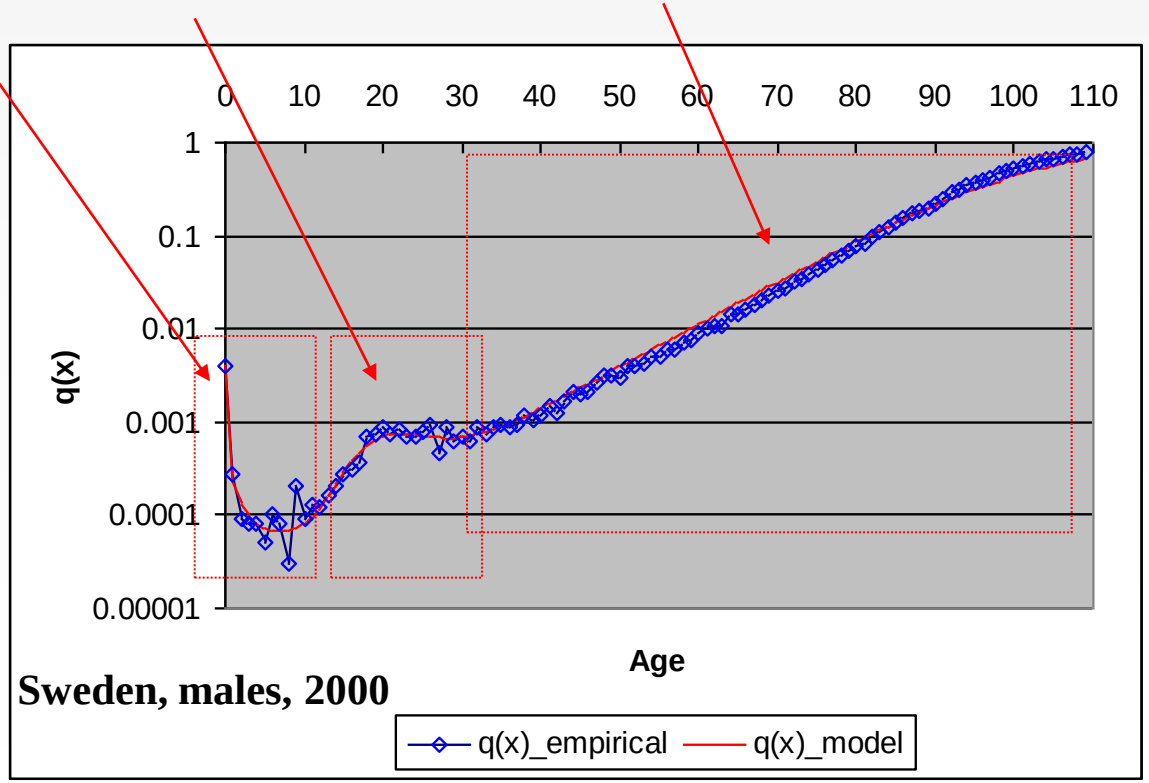
The H-P model includes three additive terms and eight parameters A, B, C, D, E, F, G, and H

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + D \cdot e^{-E(\ln x - \ln F)^2} + GH^x$$

The first term determines the rapid drop of mortality during the early childhood years. *A* is close to q_1. *B* reflects closeness between q_1 and q_0. *C* determines steepness of the decrease.

The second term reflects excess mortality from external causes for males and external and maternal causes for females. *D* indicates severity, *E* reflects the breadth of the hump, *F* represents the mean age of the hump.

The third term is a modified version of the Gompertz exponential reflecting the rise of mortality with age over adult ages.



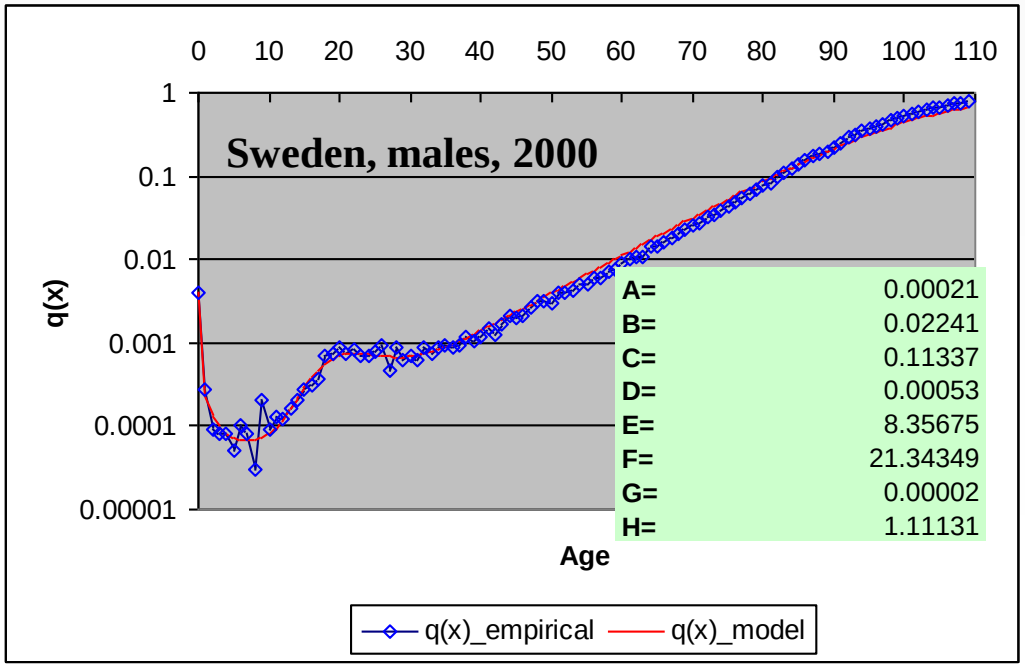
Fitting the H-P model with the Excel Problem-solver

We estimate the eight H-P parameters by solving the following optimization problem:

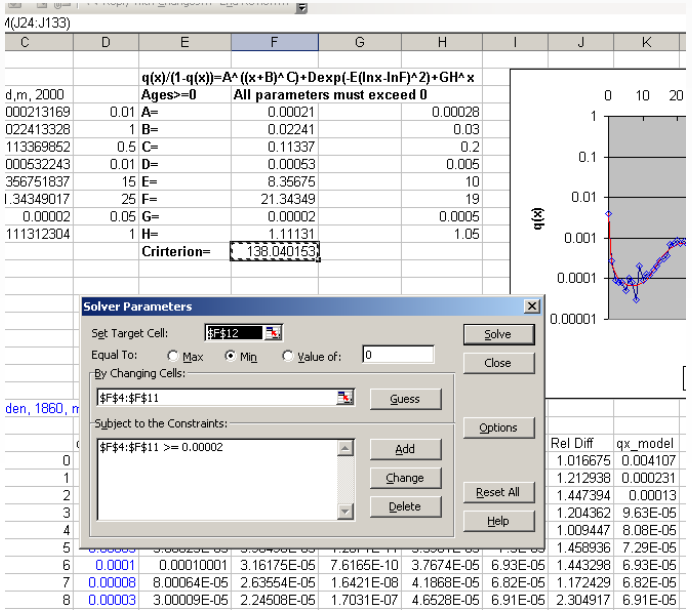
$$\sum_x \max\left(\frac{q_x/(1-q_x)}{\xi_x}, \frac{\xi_x}{q_x/(1-q_x)}\right) \rightarrow \min$$

where: $$\xi_x = A^{(x+B)^C} + D \cdot e^{-E(\ln x - \ln F)^2} + GH^x$$

with *A, B, C, D, E, F, H, G* being positive. **Problem** - instability of the estimates.

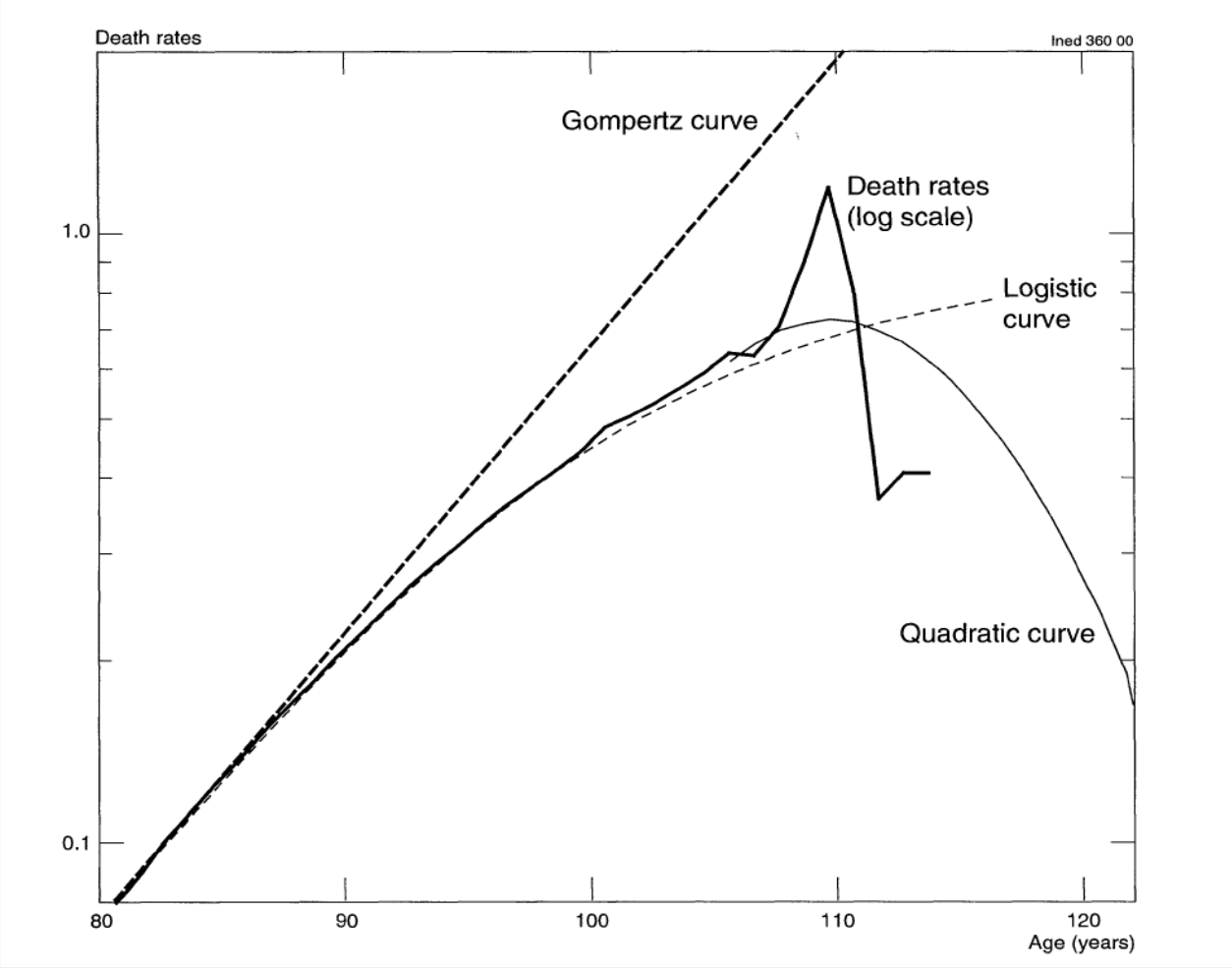


Heligman-Pollard.xls



Enigma of the oldest age. Which model applies?

Death rates in aggregation of 14 countries with good quality data (Western Europe + Japan)



Source: Vaupel et al., 1998.

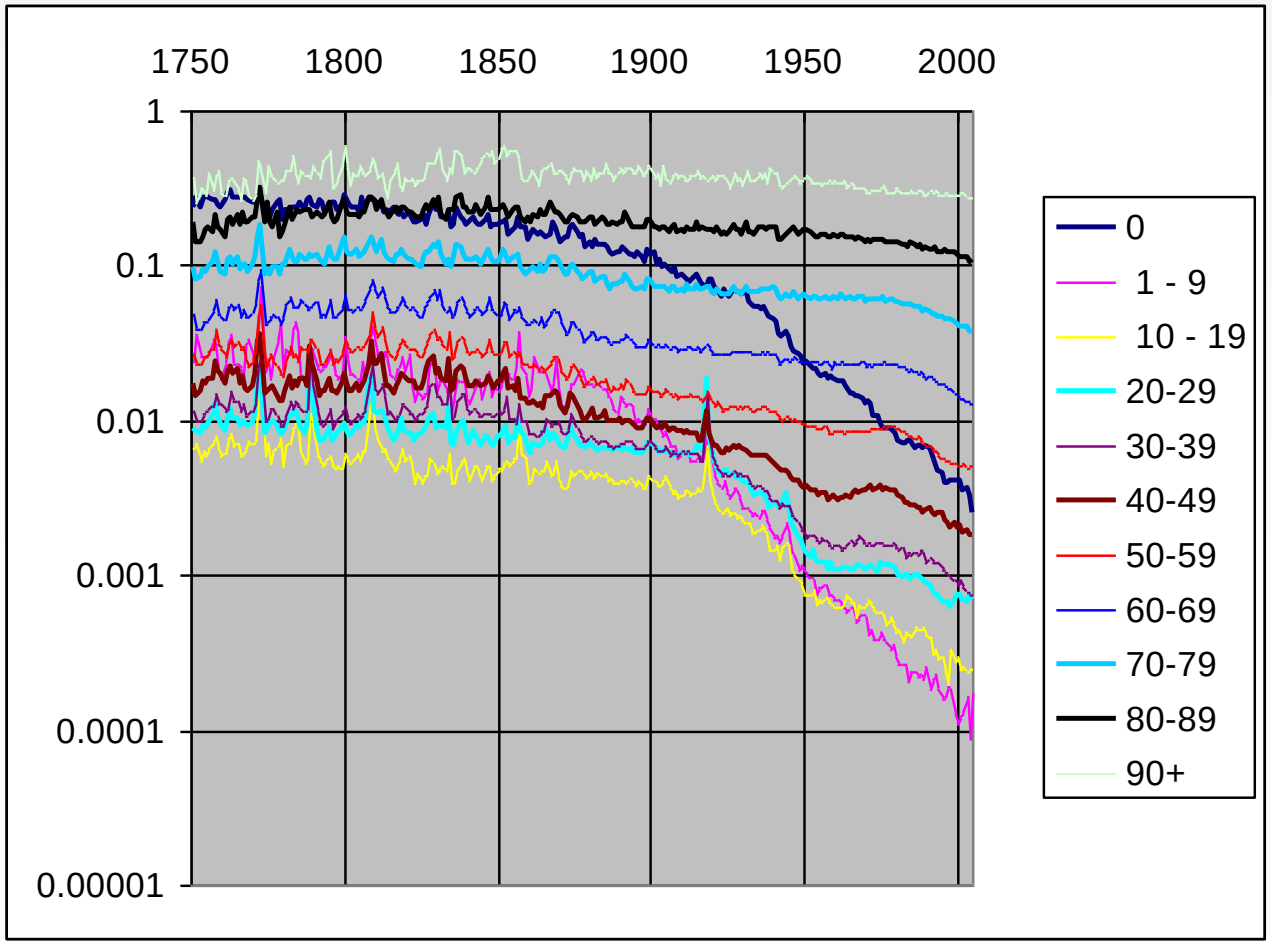
TIME. Uneven speed of temporal decreases depending on time period and age

Death rates for Sweden in 1750-2005.

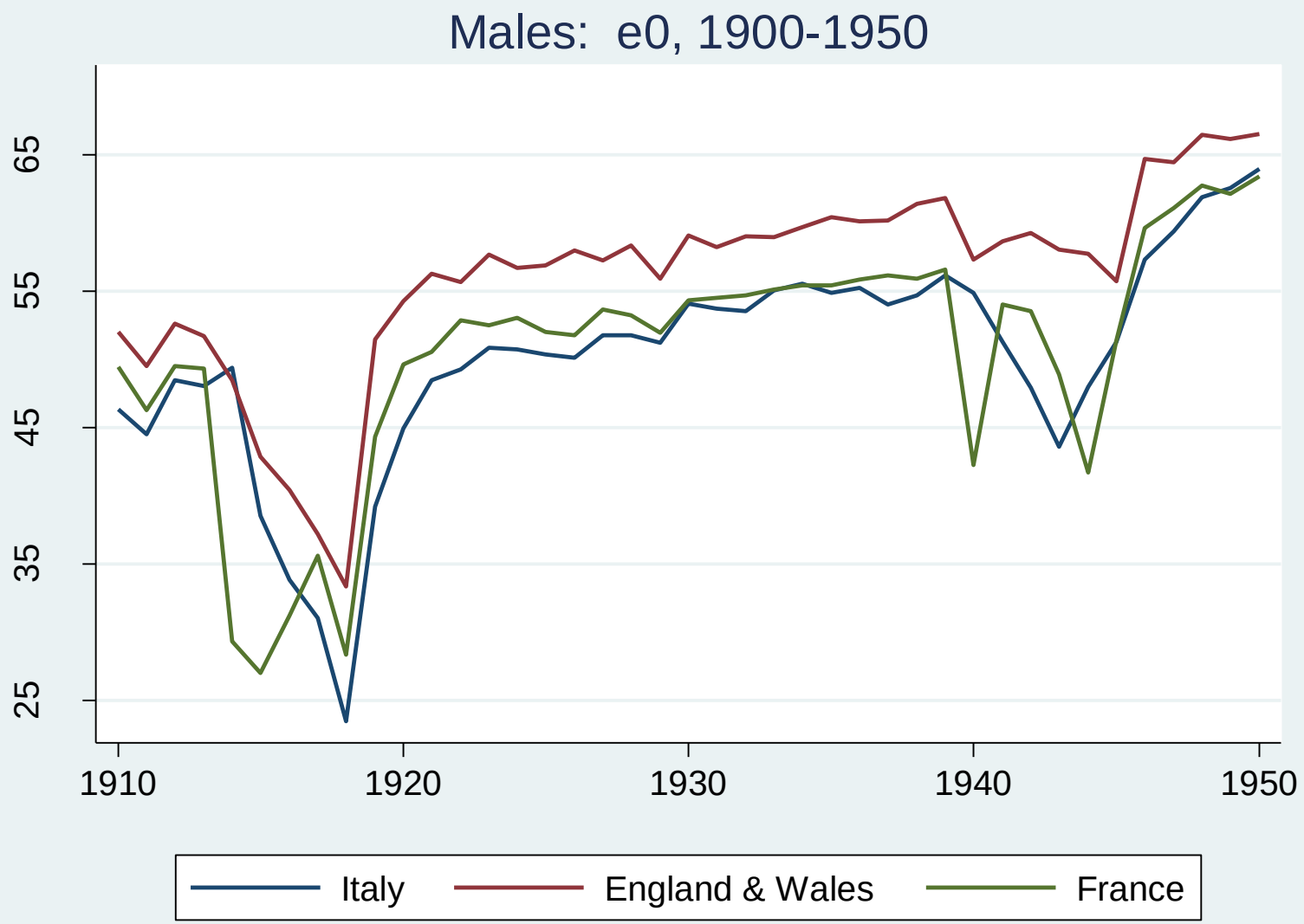
Before the end of 19th century, mortality decline was very slow. During the first half of the 20th century, mortality at young ages was rapidly decreasing. The progress at older ages was very slow.

Mortality decrease at ages over 60 has accelerated since the 1970s due to so-called “cardiovascular revolution”.

At ages 20 to 49, mortality decreased somewhat slower than at younger ages and also experienced some stagnation in the 1960s-70s.



TIME. Period effects of extreme events. Wars and epidemics



COHORT. A concept of cohort effects

Some events or exposures to certain conditions can cause an elevation or a depression of mortality over later life. In epidemiology, analyses linking today‘s health with past experiences and conditions are known as the life course analyses.

Some diseases such as stroke have been related to insufficient fetal growth.

Food shortage, vitamin or protein deficiencies in childhood can cause diseases in later life. Mortality of especially short people (a sign of food shortage in childhood) is greater than mortality of tall or normal-height people.

Poor nutrition and living conditions of children depress their immune status and make them more vulnerable in respect to infections (such as TB). Infection agents contracted in childhood can cause some deadly pathologies such as stomach cancer at adult ages.

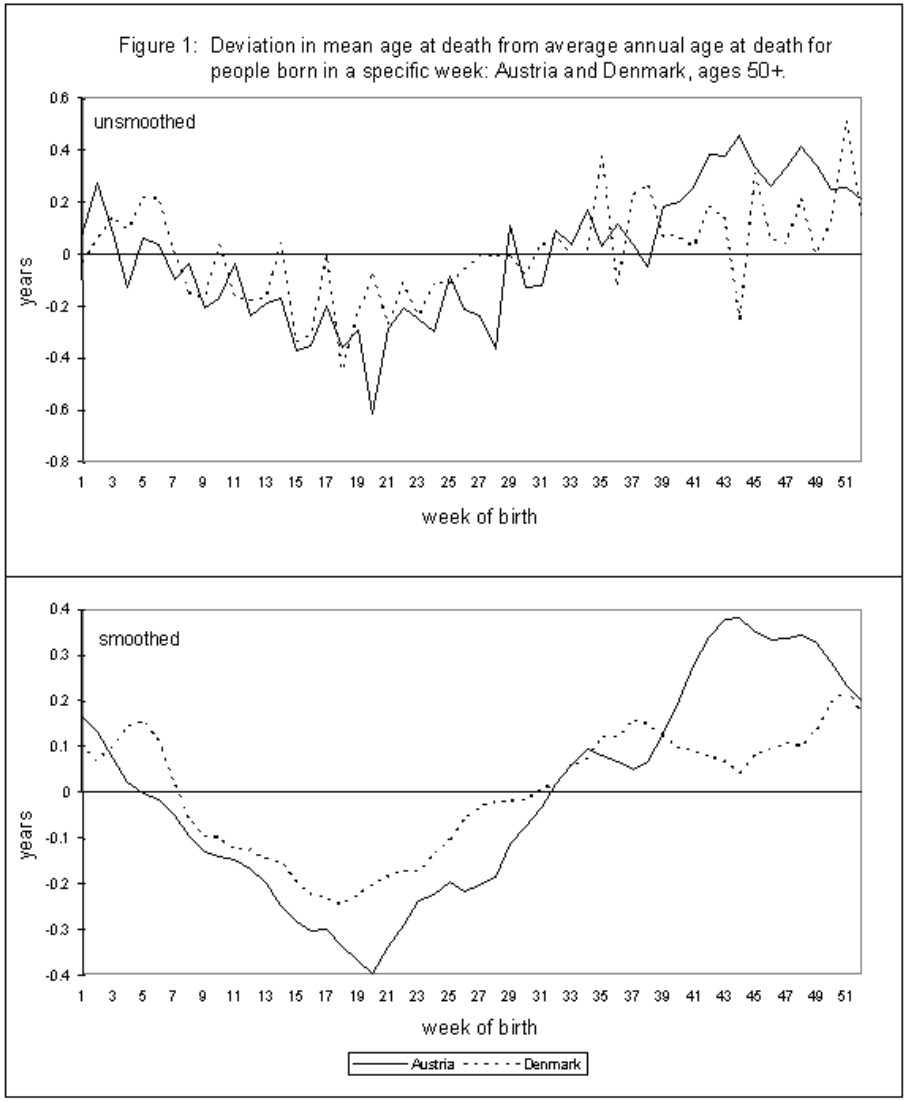
Severe stress due to participation in war operations can cause a loss of mental health and elevate risk of suicides in later life.

On the other hand, one can imagine a higher robustness of “the strongest”, who manage to survive in hard conditions.

Debilitation refers to health worsening and mortality elevation due to adverse experiences in the early life. *Selection* is an opposite phenomenon.

Example 1. Association between length of life and season of birth

Doblhammer (1999) found a mortality effect related to season of birth. After age 50, people born in winter tend to live longer than those born in summer. The effect is likely to be related to quality of mothers' nutrition during the last months of pregnancy.



Example 2: Association of mortality from stomach cancer in 1991-93 at age 65-74 with IMR in 1921-23

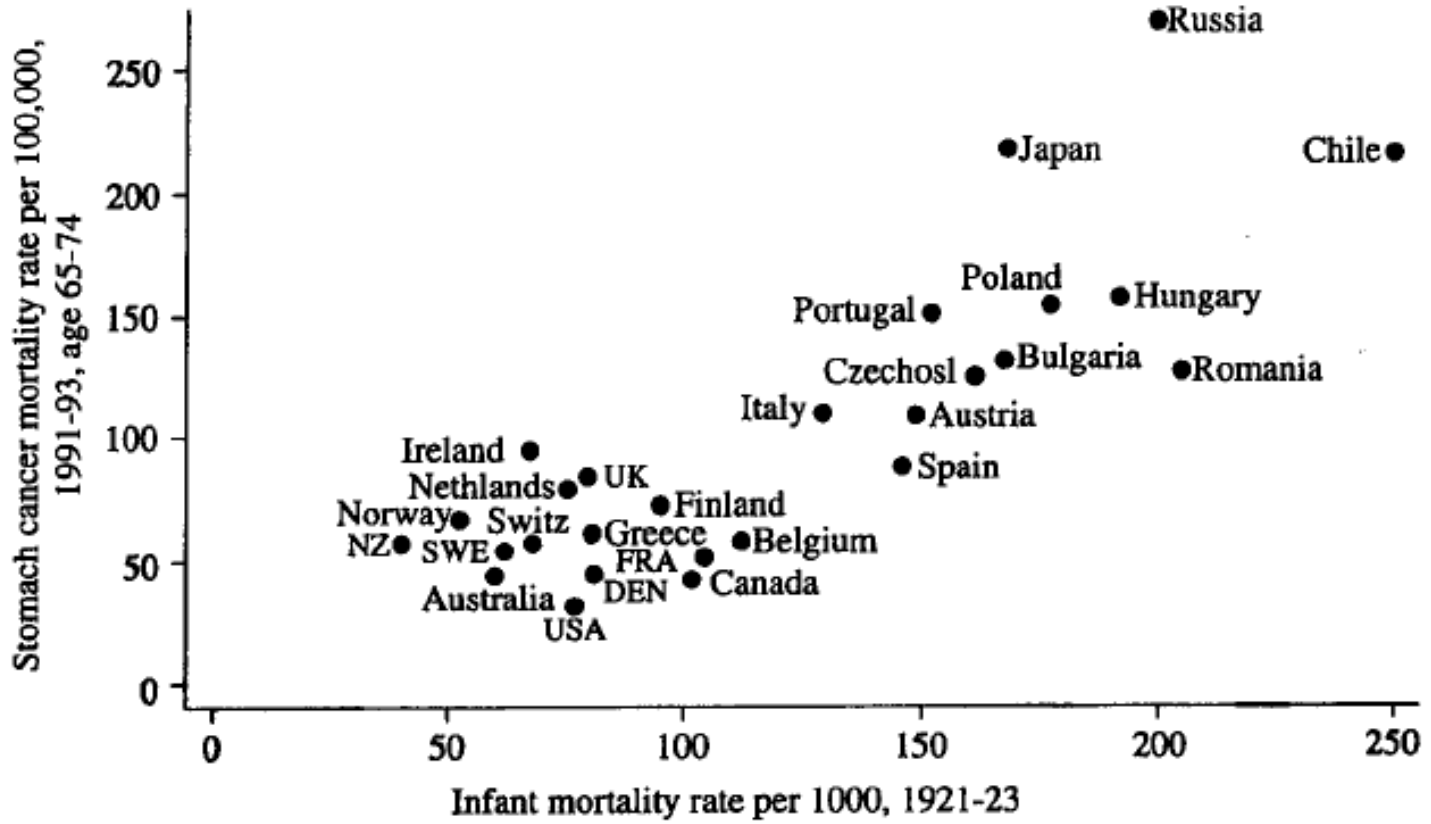
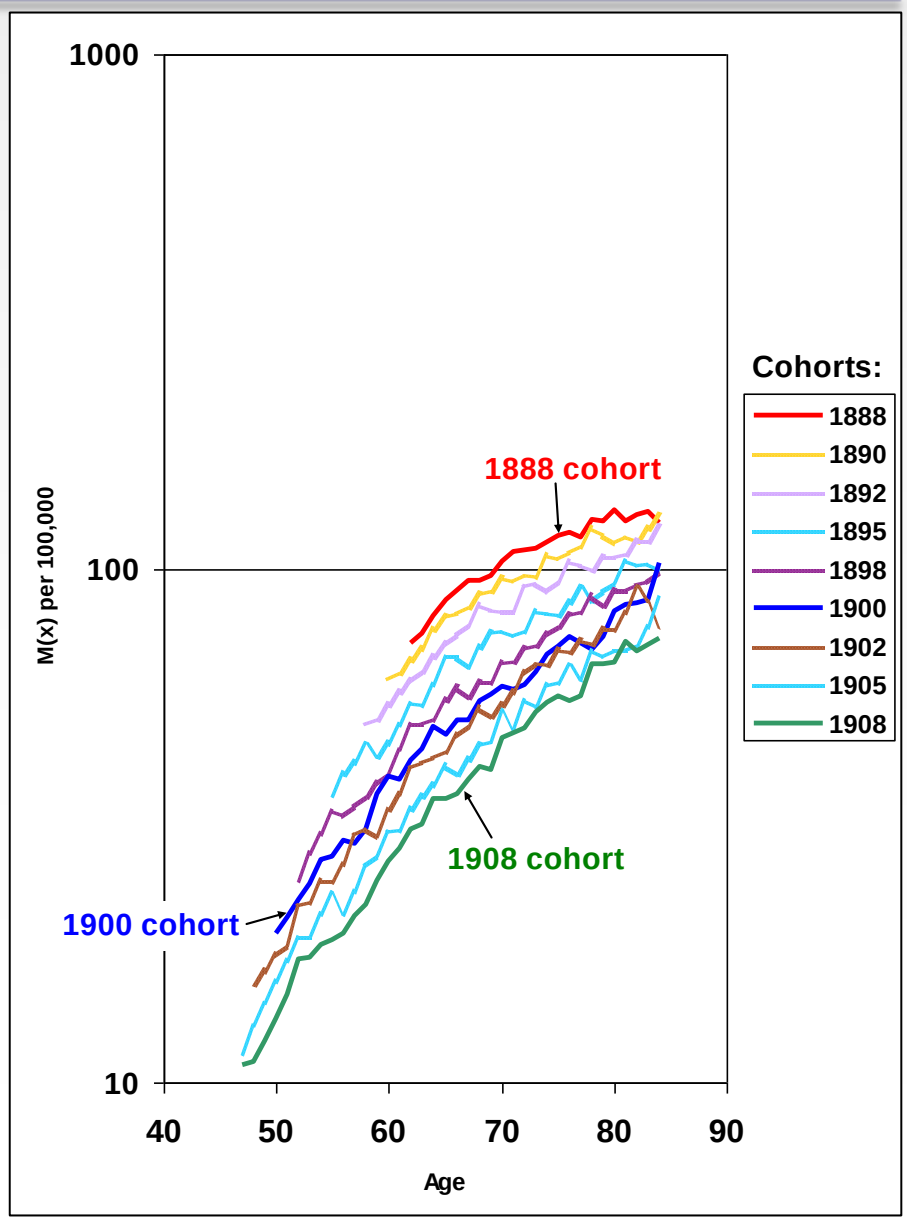


Fig. 4.14 Plot of infant mortality 1921–23 against stomach cancer mortality 1991–3 for men aged 65–74 in 27 countries
Source: Leon and Davey Smith (2000)

Example 2. Continuation: mortality from the stomach cancer by cohort and age

Age-specific cohort mortality rates for the stomach cancer.

USA males, cohorts born in 1888-1908.



Growing age-specificity in the mortality decrease across cohorts in England and Wales

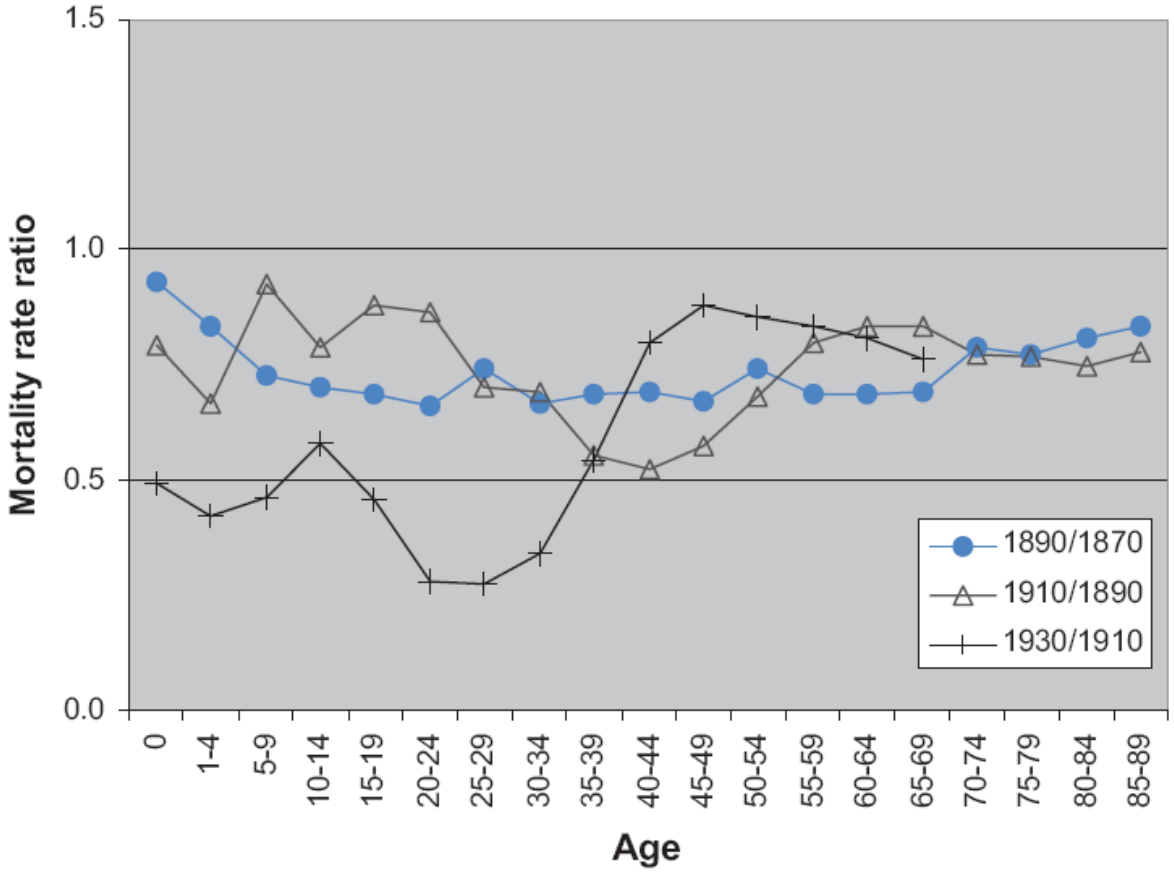


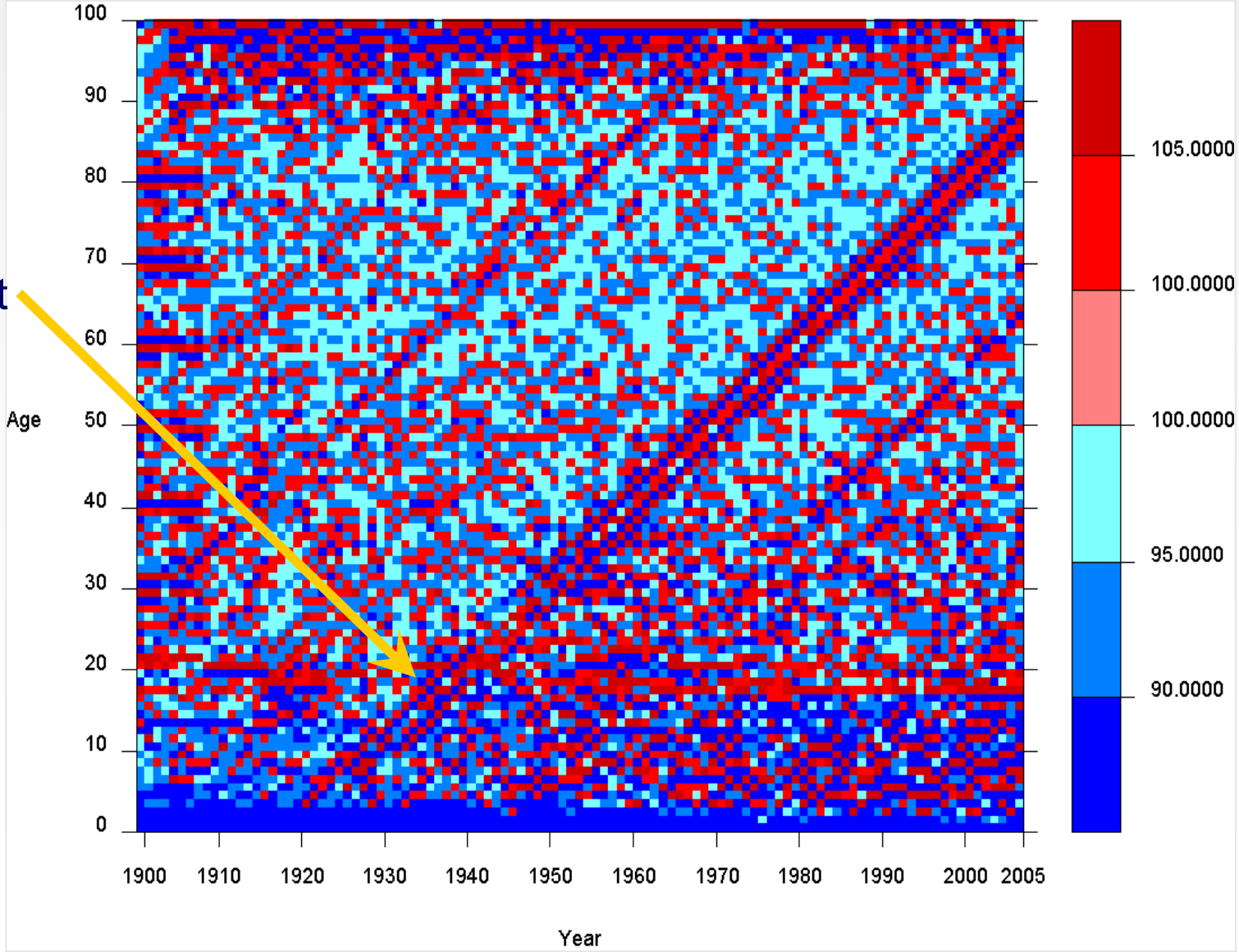
Figure 1. Age-specific ratios of mortality rates between successive female cohorts in England and Wales.

Source: Human Mortality Database (www.mortality.org).

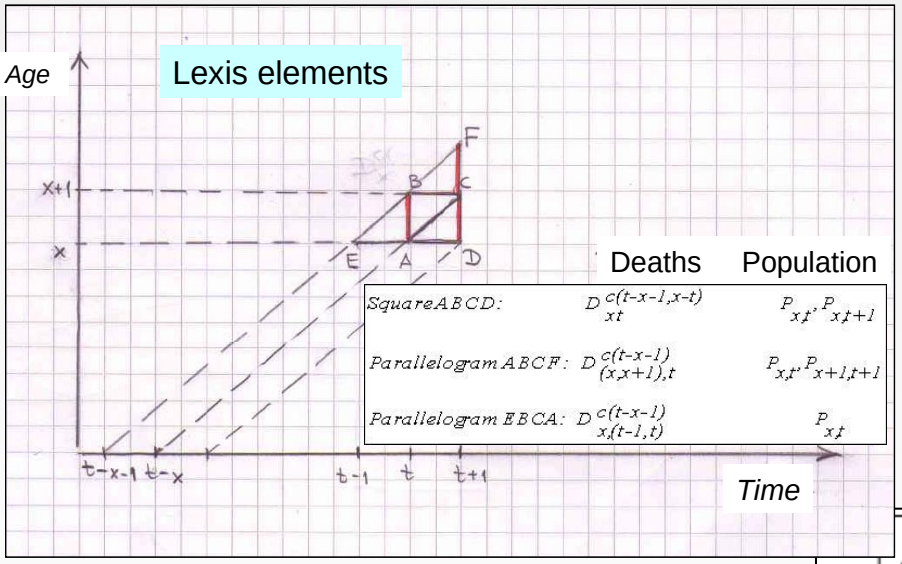
False cohort effects induced by Lexis squares. French death rates based on Lexis squares

$$K_x = M_x / (0.5(M_{x+1} + M_{x-1}))$$

Seemingly a major „cohort effect“



Mortality rates corresponding to one-year three types of Lexis elements



Type 1: $$M_{x,t} = \frac{D_{x,t}}{0.5(P_{x,t} + P_{x,t+1})}$$

Type 2: $$M^{C(t-x-1)}_{(x,x+1),t} = \frac{D^{C(t-x-1)}_{(x,x+1),t}}{0.5(P_{x,t} + P_{x+1,t+1})}$$

Type 3: $$M^{C(t-x-1)}_{x,(t-1,t)} = \frac{D^{C(t-x-1)}_{x,(t-1,t)}}{P_{x,t}}$$

Infant mortality „rate“ (death probability):

$$IMR_{0,t} = \frac{D_{0,t}}{B_t}$$

B_t – births in year t

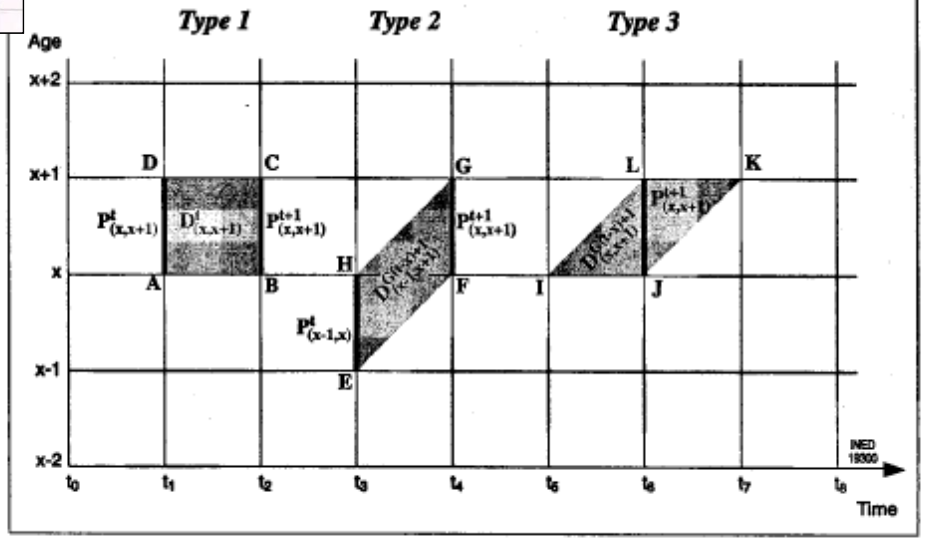
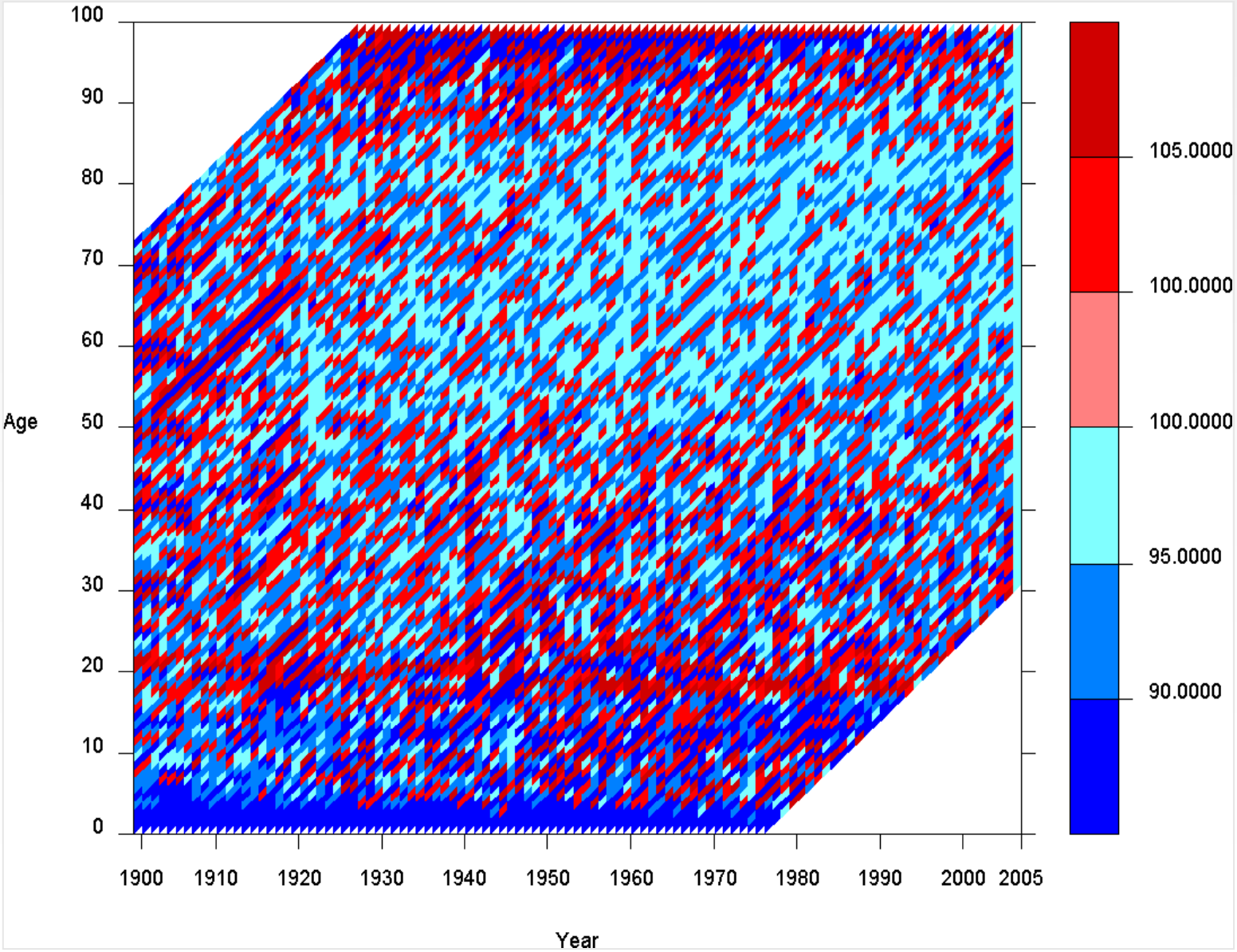


FIGURE 6–8 Representation of the elements required for computing a death rate by age according to the classification mode of the deaths.

The French death rates based on Lexis parallelograms



Task

1. Использую данные для женщин Швеции из HMD оцените параметры модели Гомперца за каждые 50 лет начиная с 1760 года.
2. Опишите каким изменения кривой дожития соответствуют изменения параметров модели

Дополнительный материал для самостоятельного изучения

APC. Age-period-cohort analysis of mortality. The identification problem.

APC is a method for objective detection of cohort effects in mortality. The central difficulty is to identify a unique set of parameters that provide optimal fit to the observed mortality data. The standard APC model has the following form:

$$f(M_{i,j,k}) = \alpha_i + \beta_j + \theta_k + \varepsilon_{i,j,k}$$

$M_{i,j,k}$ is an observed death rate corresponding to age *i*, year *j*, and cohort *k*; *f(.)* is its transformation such as log or logit; the three parameters (dummy variables) describe patterns of change in $f(M_{i,j,k})$ by age, period, and cohort, respectively; epsilon is an error term.

First, there is a simple identification problem common for most of regression models. Let us denote the three parameter estimates by „cap“ then

$$f(\hat{M}_{i,j,k}) = \hat{\alpha}_i + \hat{\beta}_j + \hat{\theta}_k$$

The predicted value of the transformed rates is equal to the sum of the estimated age, period, and cohort effects. It is clear that there are many sets of the three parameters that fulfill this equation. So, as usual in regression analyses, for certain *i,j,* and *k*, the parameter levels should be fixed at 0.

$$\hat{\alpha}_{i'} = \hat{\beta}_{j'} = \hat{\theta}_{k'} = 0.$$

Source: Wilmoth, 2006

APC. Age-period-cohort analysis of mortality. The identification problem-2.

Another type of the identification problem is specific for the APC analysis and is more challenging. Let us make a linear transformation of the three estimated parameters:

$$\hat{\alpha}^*_i = \hat{\alpha}_i + \lambda \cdot i \ , \quad \hat{\beta}^*_j = \hat{\beta}_j + \lambda \cdot j \ , \quad \hat{\theta}^*_k = \hat{\theta}_k - \lambda \cdot k \ .$$

Obviously, with any lamda value, the sum of the three effects will remain unchanged because *k=j-i*.

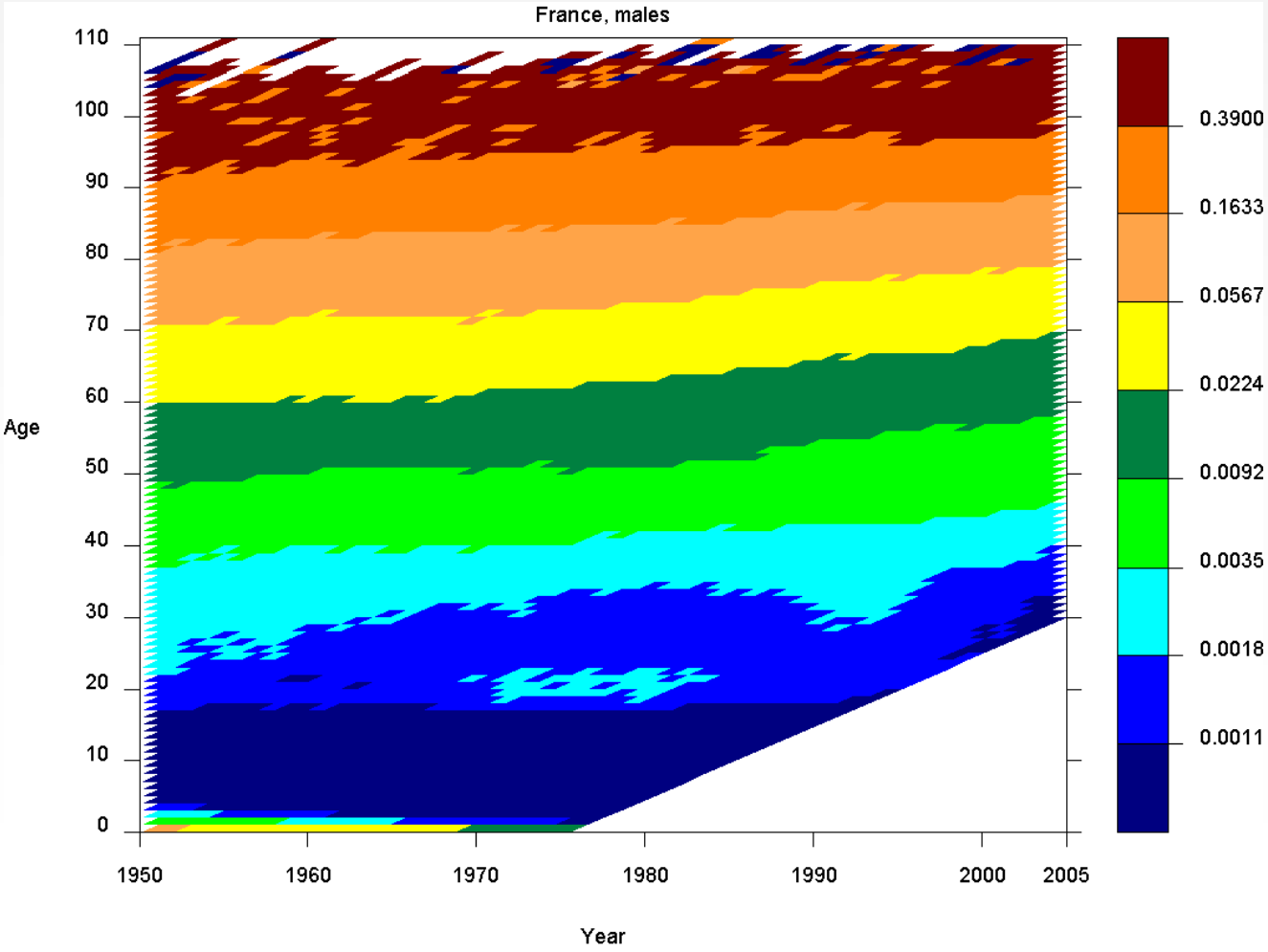
$$\hat{\alpha}^*_i + \hat{\beta}^*_j + \hat{\theta}^*_k = \hat{\alpha}_i + \hat{\beta}_j + \hat{\theta}_k .$$

To overcome the identification problem, a proper strategy is to avoid including all three sets of parameters in the model and to focus on two dimensions only treating the third dimension as a sort of residual.

Source: Wilmoth, 2006

APC. Age-period-cohort analysis of the French mortality in 1950-2005. Data.

Input data array: $M_{i,j,k}$ calculated from elementary Lexis parallelograms with horizontal sides. Filename: *apc_france-females.dta*



APC. Age-period-cohort analysis of the French mortality in 1950-2005. Method.

Input data array: $M_{i,j,k}$ calculated from elementary Lexis parallelograms with horizontal sides. Filename: *apc_france-females.dta*

The model:

$$\ln(M_{i,j,k}) = \text{age} + \text{period} + \text{age0\_15} \times \text{period} + \text{age16\_35} \times \text{period} + \text{cohort}$$

The two interaction effects allow to capture a steeper pace of mortality decrease at ages 0 to 15 and a slower decrease at ages from 16 to 35 years.

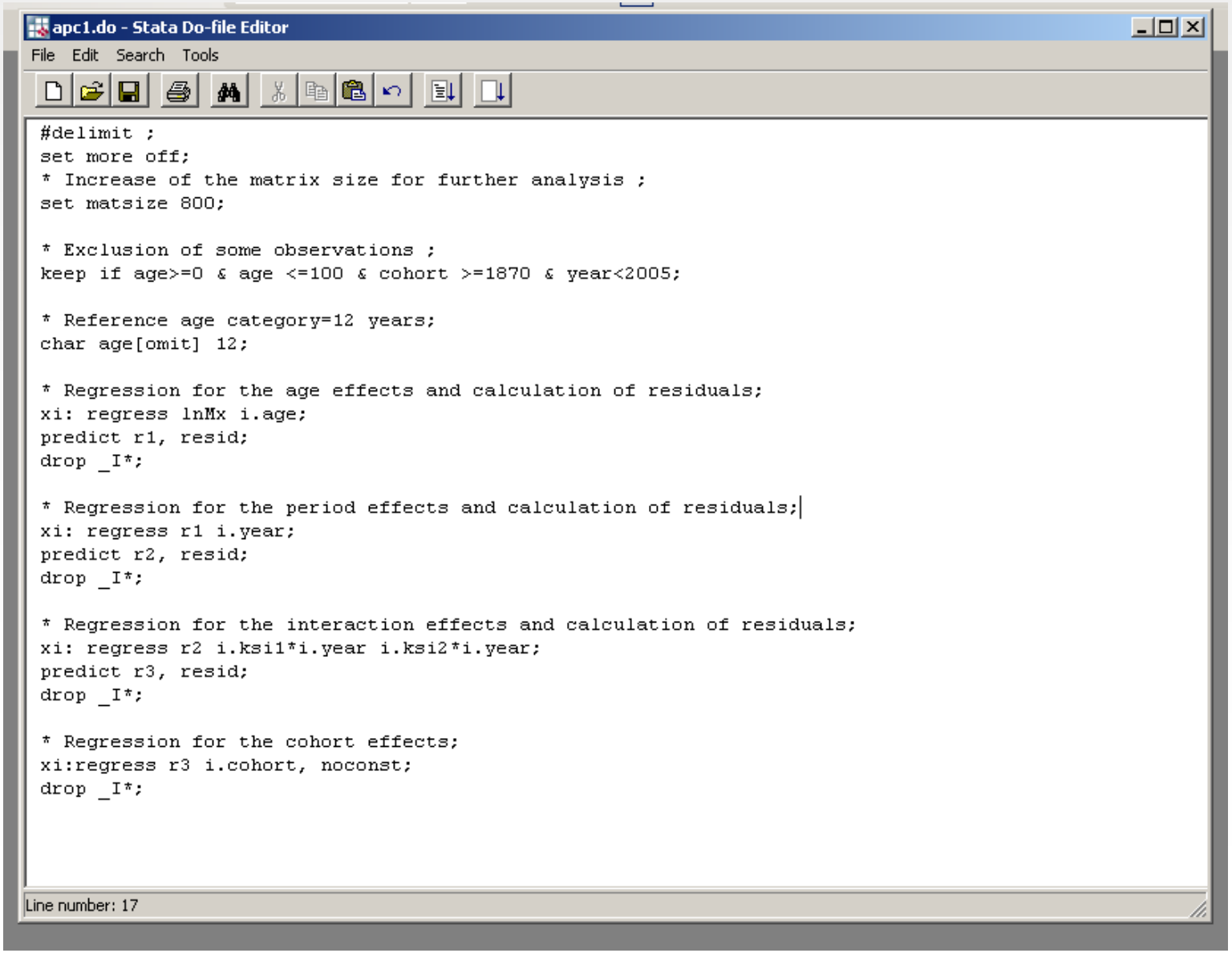
Step 1: OLS regression of *ln(M)* on age $\Rightarrow$ residual1

Step 2: OLS regression of residual1 on period $\Rightarrow$ residual2

Step 3: OLS regression of residual2 on the two interactions $\Rightarrow$ residual3

Step 4: OLS regression of residual3 on cohort

APC. Age-period-cohort analysis of the French mortality in 1950-2005. Stata code.



APC. Age-period-cohort analysis of the French mortality in 1950-2005. Results.