# Chapter 5
# Surface Plots of Smoothed Mortality Data

## 5.1  From Raw Death Rates to Smoothed Death Rates

We have seen in the previous chapter, that "raw" death rates can suffer from considerable random fluctuations. Assuming that data quality is not an issue, this noise can be caused by (1) very few numbers of deaths (numerator), by (2) very few persons exposed to the risk of dying (denominator) or by (3) small populations in general. Problem (1) typically occurs at young ages. We selected age 15 in France in Panel (a) of Fig. 5.1. Despite a large population in general, deaths occur—thankfully—relatively rarely at that age. (2) The opposite is true at advanced ages as shown in the middle panel of the same figure. Very few people are still alive at age 95 in Italy, although it is a large population having relatively high life expectancy. Problems (1) and (2) occur in countries with tens of millions of people only at young and old ages. The smaller the population size, the more ages are affected. Panel (c) illustrates issue (3) using Danish data. The mortality trajectory in highly developed countries is rather smooth around age 80. In countries with just a few millions of people, considerable random fluctuations can be even observed there. Please note that more than five million people live in Denmark. Hence, the challenge becomes even bigger in smaller countries such as the Baltic states, Luxembourg or, especially, in Iceland.

We decided therefore to smooth the data. Myriads of methods exist to smooth data. While the pattern over age can be appropriately captured by parametric models, the trajectory over time differs considerably between ages and countries. Our decision was therefore to use a non-parametric smoothing approach. We selected the so-called *P*-spline approach, originally developed by Eilers and Marx (1996), adapted to the analysis of mortality by Currie et al. (2004) and further refined by Camarda (2008). The author, Carlo Giovanni Camarda, also provides the R extension package "MortalitySmooth" (Camarda 2012), which makes it easy and straightforward to apply the method. At its core, the model assumes
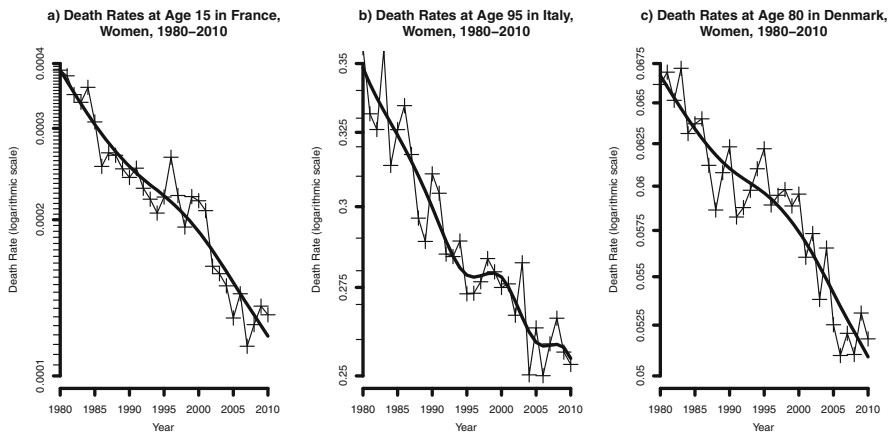
**Fig. 5.1** The necessity to smooth raw death rates. Using data for France, Italy, and Denmark, panel (**a**), (**b**) and (**c**) illustrate three sources of random fluctuations: few numbers in the numerator (Panel (**a**) for age 15), few numbers in the denominator (Panel (**b**) for age 95) or small population sizes in general (Panel (**c**) for age 80) (Data source: Human Mortality Database)

Poisson distributed death counts with the (log-)exposures as an offset to account for changing population sizes over time and/or age. The method uses *B*-splines as regression bases. Whereas the number and position of the basis functions is crucial for standard smoothing with *B*-splines, the *P*-spline approach uses "too many" bases, which would normally result in overfitting. The *P* in the name of the method refers to the penalization of adjacent regression coefficients that differ too much from each other. Further technical details about the basis functions, the order of the differences, the penalty term $\lambda$, etc. are extensively discussed in the aforementioned references. The bold solid black lines in each panel of Fig. 5.1 depict the data smoothed with *P*-splines for the three given ages over time. One can easily recognize that the selected smoothing method is flexible enough to model irregular developments but is not prone to overfit the data.

The univariate time series of Fig. 5.1 is synthetic. Only cartoon characters such as Bart Simpson or Eric Cartman can retain their age over time. In reality, each individual is 1 year later 1 year older. Therefore we smoothed the data simultaneously over age and time using the function `Mort2Dsmooth` of Camarda's package "MortalitySmooth" (2012).

Raw death rates for Estonian women aged 60–80 years from 1980 to 2000 are illustrated in the left panel of Fig. 5.2 as a three-dimensional mortality surface. The general shape of increasing mortality over age can easily be observed. The right panel, featuring smoothed data, also shows the decline in mortality at higher ages over time, which is difficult to track down in the presence of noise in the data.
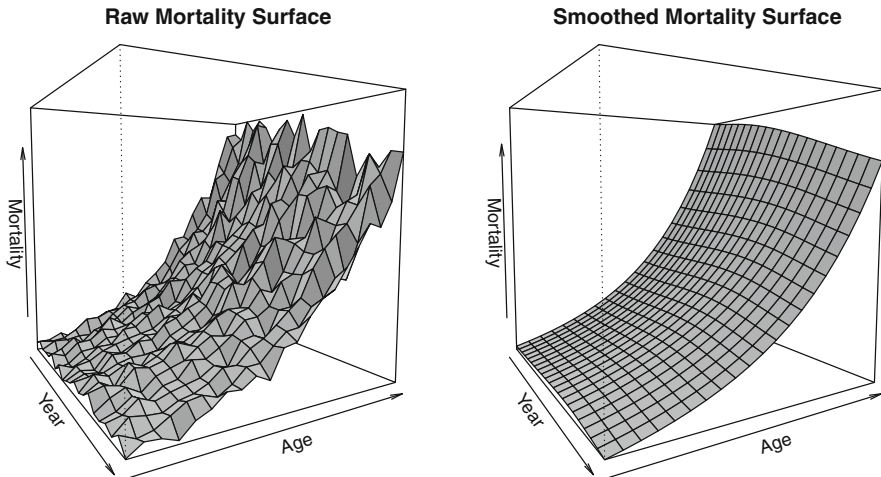
**Fig. 5.2** 3D plot of raw and smoothed death rates of Estonian women aged 60–80 years in 1980–2000 (Data source: Human Mortality Database)

The selected three-dimensional perspective plot appears appealing at first sight. The choice of angle and elevation is somehow arbitrary, though, and allows to accentuate certain features and suppress others. Since we often want to use the mortality surface for exploratory purposes, we have to give equal exposure to each unit. Therefore, we projected the three-dimensional data on the two-dimensional Lexis-plane, denoting the level of mortality by different colors (see Fig. 5.3 as an example).

Comparable to topographic maps, we added contour lines to depict the same levels of mortality. The general upward tendency of the contour lines indicate that the same level of mortality is shifting to higher and higher ages. Thus, for a given age mortality is decreasing, resulting in an increase in life expectancy.

## 5.2  Results

Figures 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, and 5.11 depict the same set of countries as Figs. 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, and 4.8 in Chap. 4 for a proper comparison between "raw" rates and smoothed rates.[1] The smoothed surface maps make the major trends in the data more pronounced such as almost parallel straight upward

---

[1]The appendix contains therefore also maps of smoothed death rates for France, England & Wales, and Norway. They can be found in Figs. A.7, A.8, A.9, A.10, A.11, and A.12.
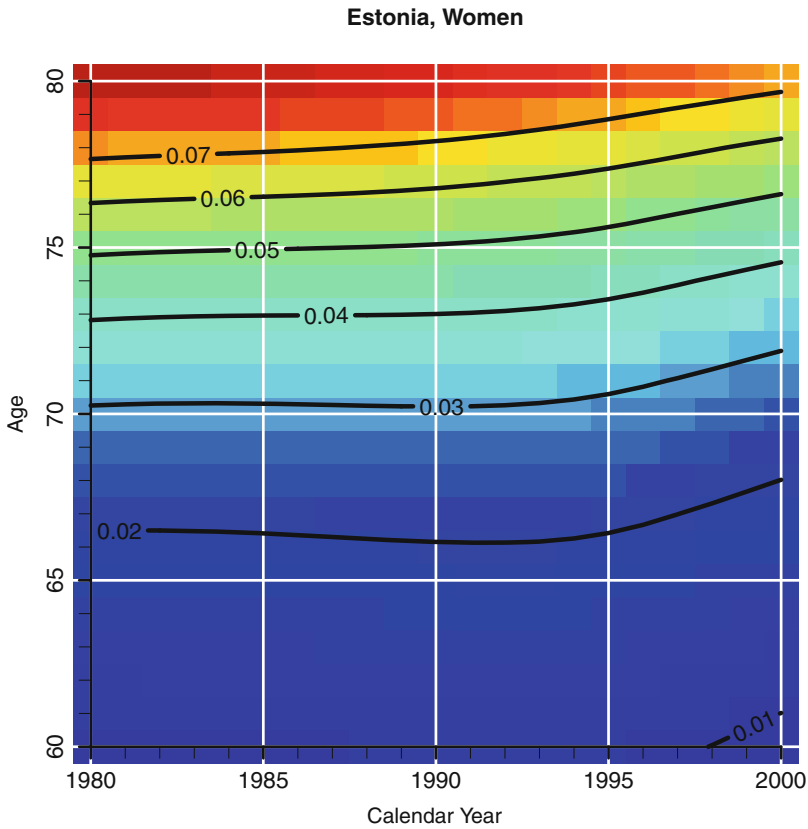
**Estonia, Women**



**Fig. 5.3** Death rates of Estonian women aged 60–80 years in 1980–2000 as an example of smoothed death rates on the Lexis plane (Data source: Human Mortality Database)

lines in Australia, Spain, and Switzerland or the sudden survival improvements in survival among young Spanish men, starting in about 1990. Also large random fluctuations due to very few deaths as we have seen in the plot of raw death rates among children in Switzerland (Figs. 4.5 and 4.6) are removed by the smoothing procedure. While smoothing intrinsically involves some dampening of sudden changes in trends, the automatic procedure to find the optimal penalizing $\lambda$s still
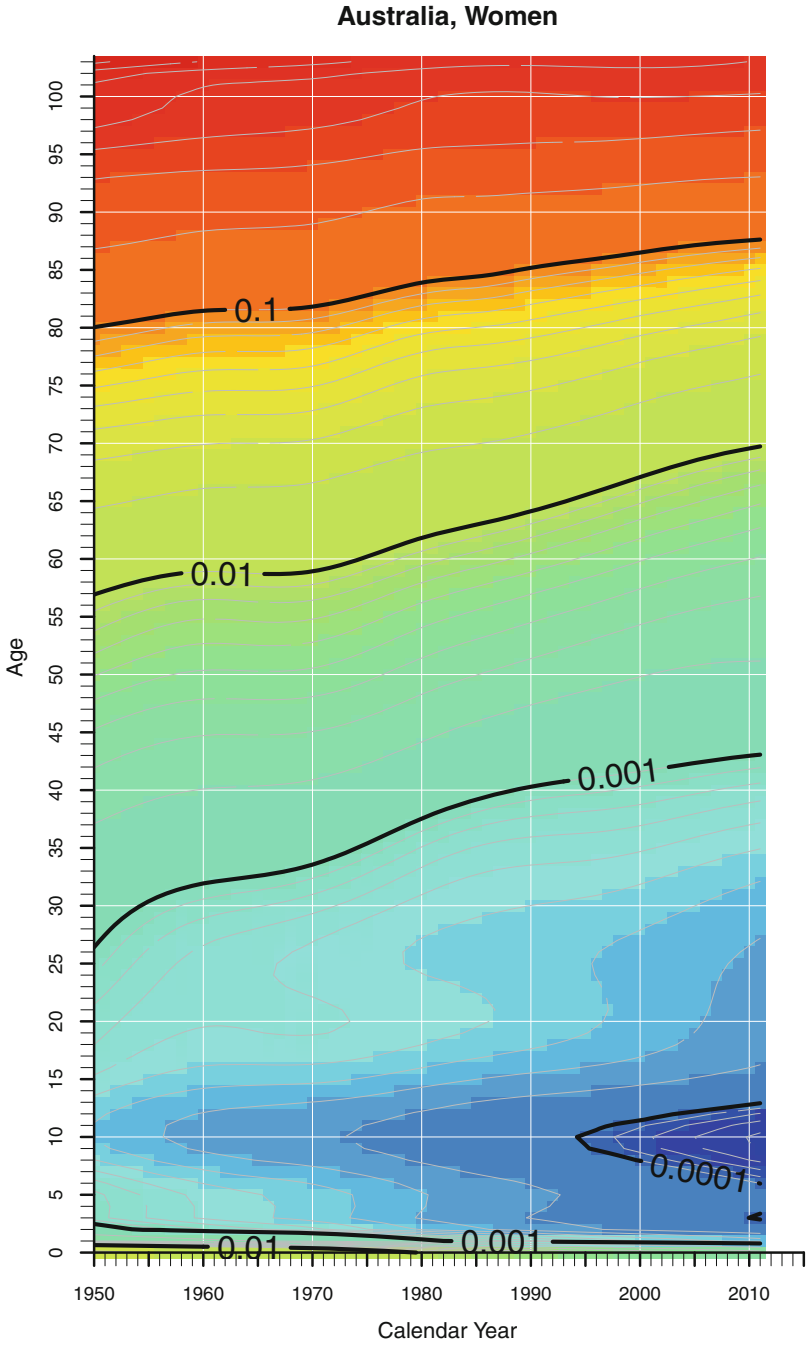
**Fig. 5.4** Smoothed death rates for women in Australia, 1950–2011 (Data source: Human Mortality Database)
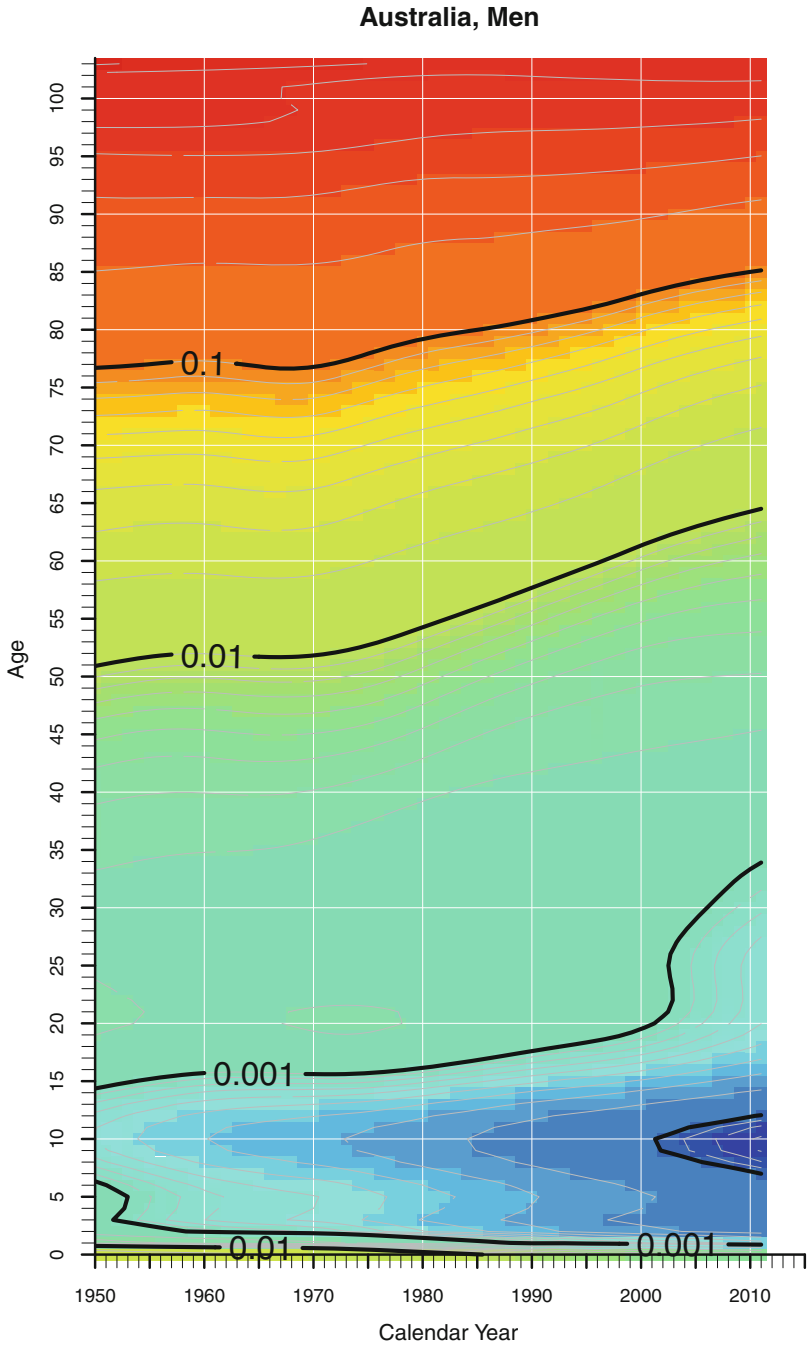
**Australia, Men**



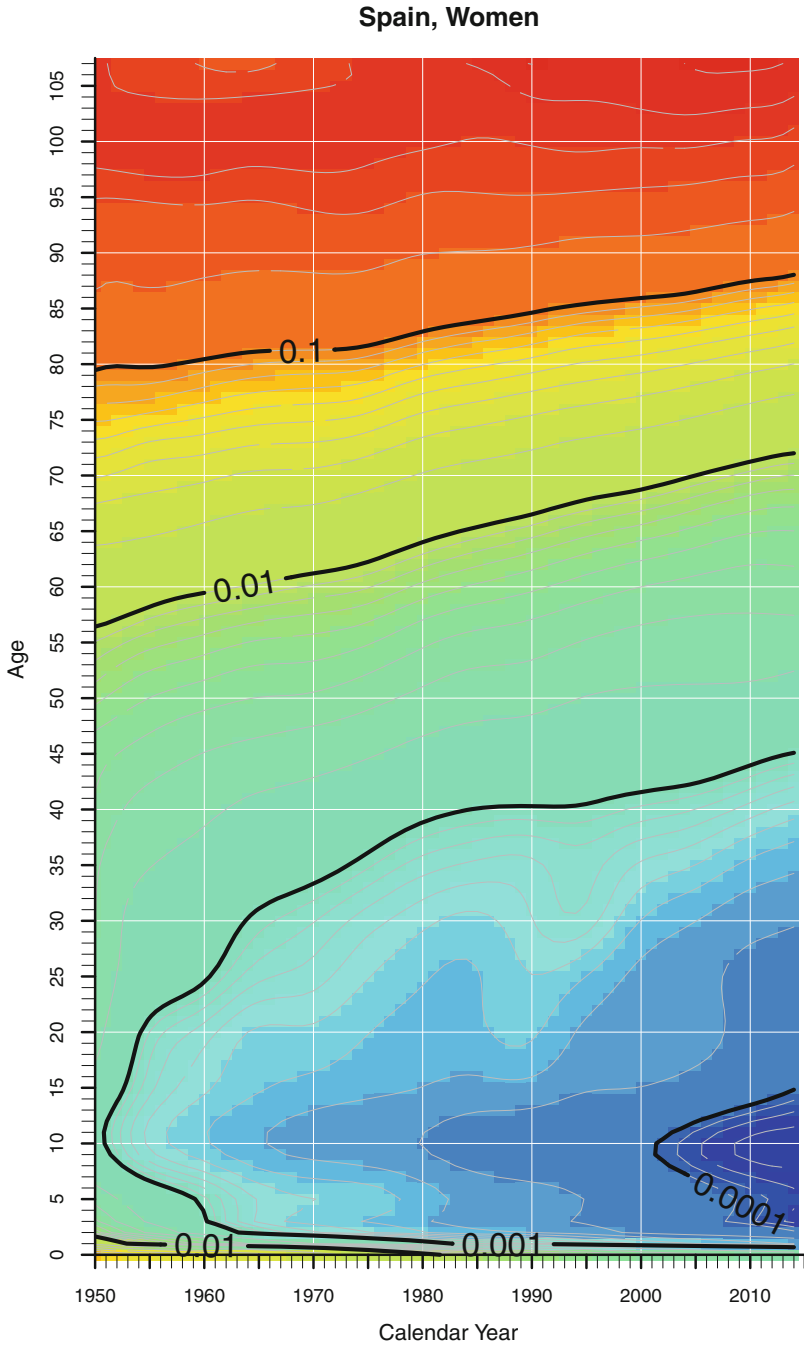**Fig. 5.5** Smoothed death rates for men in Australia, 1950–2011 (Data source: Human Mortality Database)
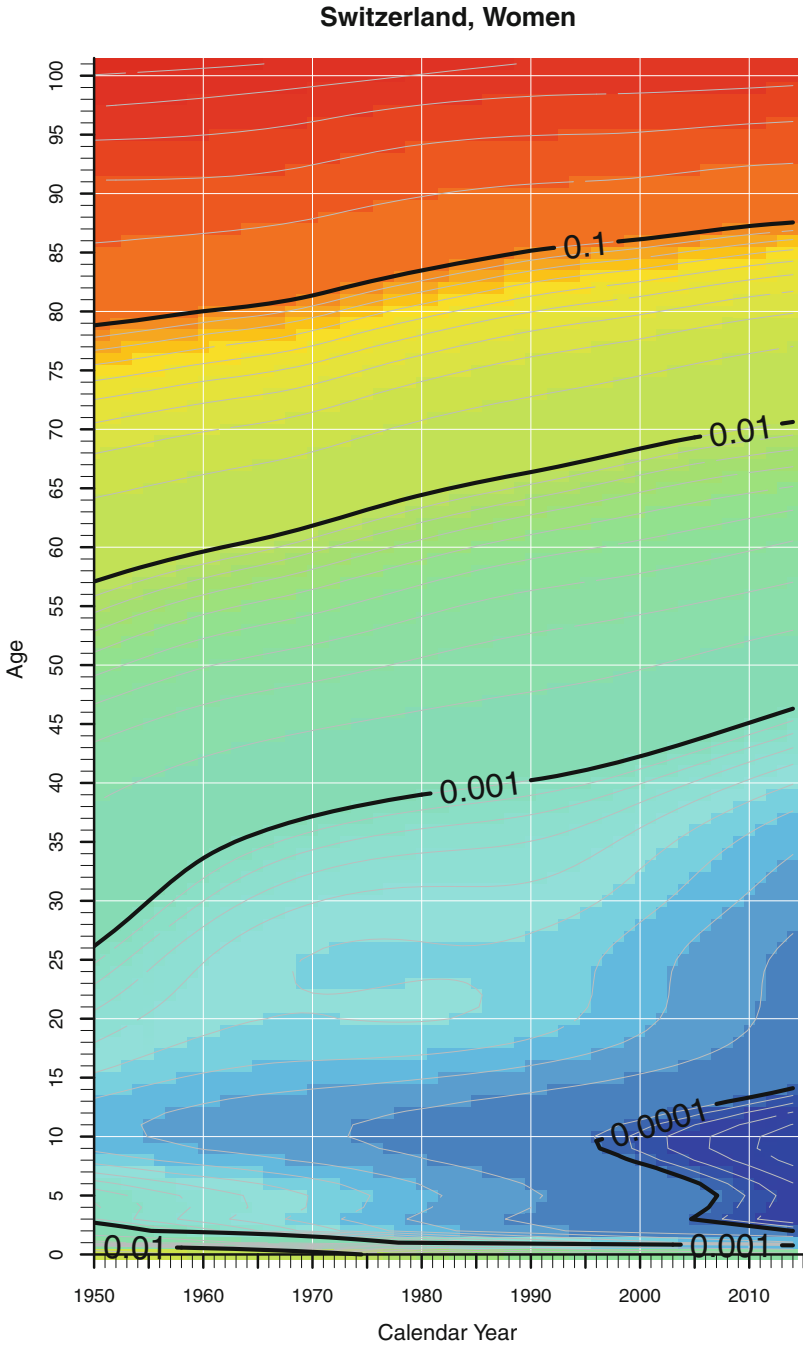
**Fig. 5.6** Smoothed death rates for women in Switzerland, 1950–2014 (Data source: Human Mortality Database)
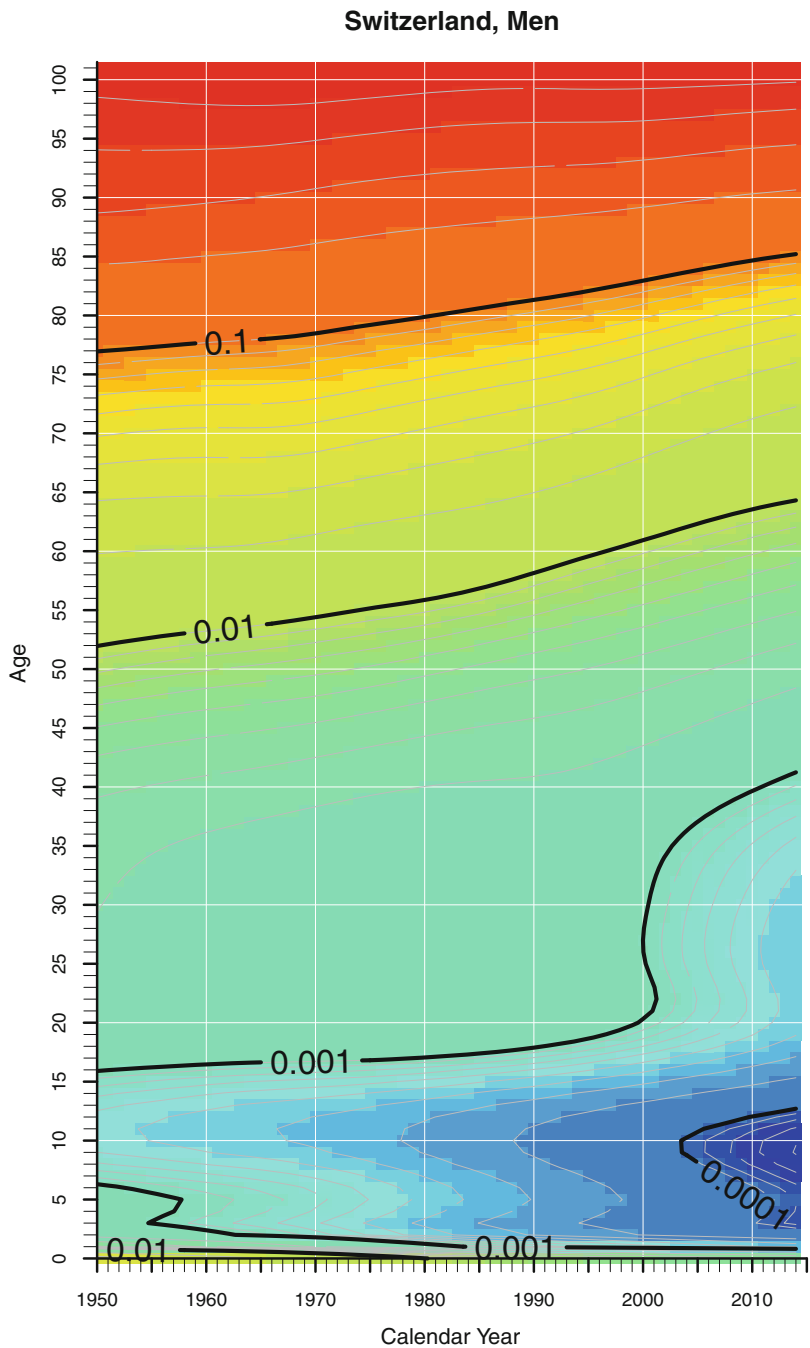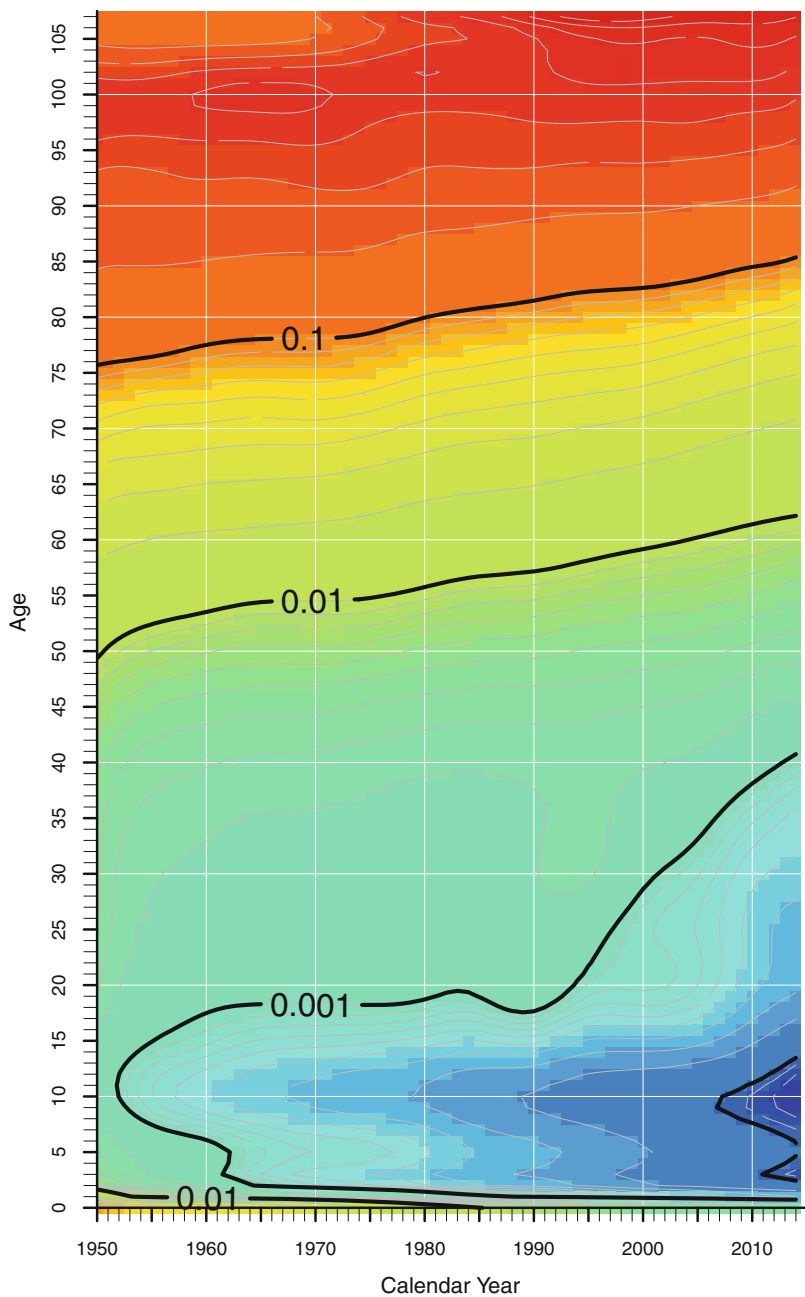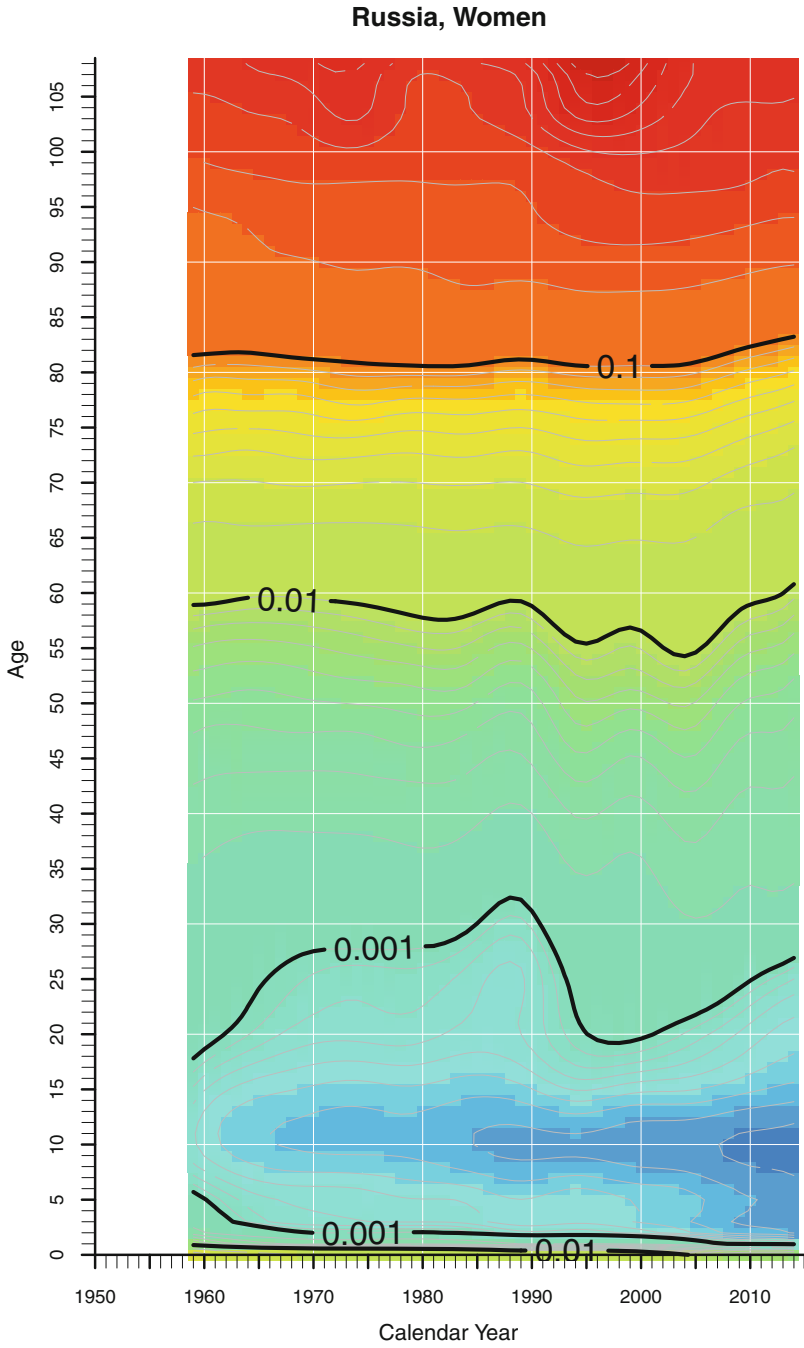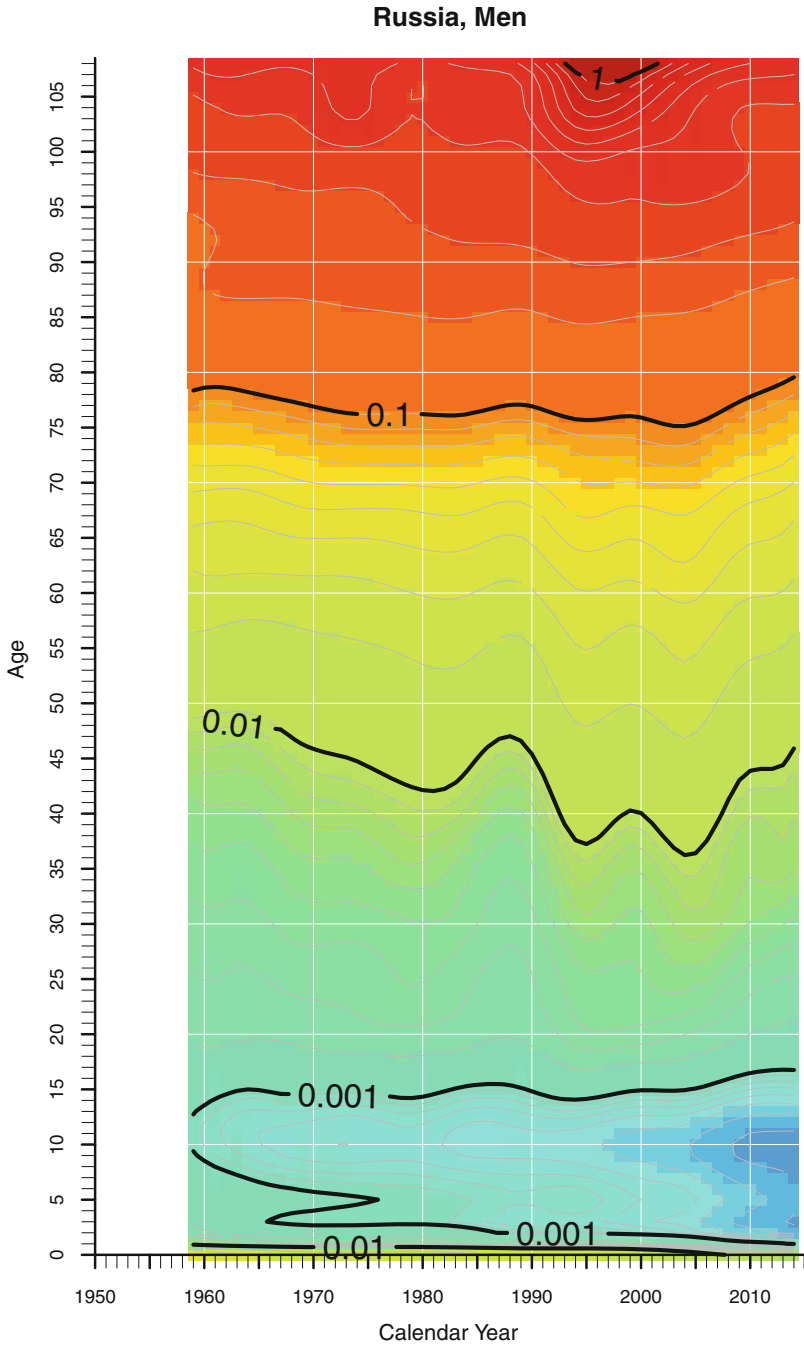
**Fig. 5.7** Smoothed death rates for men in Switzerland, 1950–2014 (Data source: Human Mortality Database)

**Fig. 5.8** Smoothed death rates for women in Spain, 1950–2014 (Data source: Human Mortality Database)

**Spain, Men**



**Fig. 5.9** Smoothed death rates for men in Spain, 1950–2014 (Data source: Human Mortality Database)

**Fig. 5.10** Smoothed death rates for women in Russia, 1959–2014 (Data source: Human Mortality Database)

**Fig. 5.11** Smoothed death rates for men in Russia, 1959–2014 (Data source: Human Mortality Database)

feature, for instance, the mortality crises among Russian men during the 1980s and 1990s. We do not want to go into further detail here as these smoothed surface maps serve as the major building blocks for the surface maps of rates of mortality improvement, which are the focus of our book and are presented in the next chapter.