## Appendix 1: Best linear estimation of $k$ (analytical solutions)

Given a set of cumulative probabilities of dying $q_i(x)$ and the level of mortality $h_i = \ln[q_i(5y)]$ of a population $i$:

$$\ln[q_i(x)] = \mathrm{a}_x + \mathrm{b}_x \cdot h_i + \mathrm{c}_x \cdot {h_i}^2 + e_i(x), \tag{A1}$$

$e_i(x)$ represents the residual at the specific age $x$, using the coefficients of the log-quadratic model $\{\mathrm{a}_x, \mathrm{b}_x, \mathrm{c}_x\}$, as shown by equation A1. These residuals from the average pattern of mortality can be defined as a function of the shape coefficient $\mathrm{v}_x$ and the scale parameter $k_i$ –which is specific to the population $i$, as shown by equation A2:

$$e_i(x) = \mathrm{v}_x \cdot k_i + \epsilon_i(x). \tag{A2}$$

The optimal value of $k_i$ minimizes the Mean Squared Error (MSE) of $\epsilon_i$, as a weighted function responding to unequal age intervals. In particular, the function $w(x)$ is assumed to be proportional to the last age interval before the age $x$, weighting the marginal contribution of an additional equation in the model.

$$MSE_i = \sum_{x \in X} w(x) \cdot \epsilon_i(x)^2. \tag{A3}$$

As a least squares' solution, the MSE is minimized by making the first derivative of equation A3 with respect to $k_i$ equal to zero, as shown:

$$\frac{\partial MSE_i}{\partial k_i}: -2 \cdot \sum_{x \in X} w(x) \cdot e_i(x) \cdot \mathrm{v}_x + 2 \cdot k_i^* \cdot \sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2 = 0.$$

The resulting estimator of $k_i$ is given by equation A4:

$$k_i^* = \frac{\sum_{x \in X} w(x) \cdot e_i(x) \cdot \mathrm{v}_x}{\sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2}. \tag{A4}$$

The uncertainty of the model is measured by the variance of the estimator. The first step is to contrast the estimated value of $k_i^*$ with the expected value of $k_i$, given that:

$$k_i^* = k_i + \frac{\sum_{x \in X} w(x) \cdot \epsilon_i(x) \cdot \mathrm{v}_x}{\sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2}.$$

Hence, the estimated variance of $k_i^*$ is defined as:

$$\mathrm{Var}[k_i^*] = \frac{\mathrm{E}\left[\left(\sum_{x \in X} w(x) \cdot \epsilon_i(x) \cdot \mathrm{v}_x\right)^2\right]}{\left(\sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2\right)^2}.$$

Assuming that prediction errors of different ages are not correlated, the expected value of $\epsilon_i(x) \cdot \epsilon_i(y)$ is zero for each age $x \neq y$. Hence, the variance of the estimator can be simplified to be:

$$\mathrm{Var}[k_i^*] = \frac{\sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2 \cdot \mathrm{E}\left[w(x) \cdot \epsilon_i(x)^2\right]}{\left(\sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2\right)^2}.$$

Under the assumption of homoscedastic errors, $\mathrm{E}[w(x) \cdot \epsilon_i(x)^2] = {\sigma_i}^2$ for all $x$; and the variance of the estimator is a function of the variance of the error of prediction $\epsilon_i$, to the form:

$$\mathrm{Var}[k_i^*] = \frac{{\sigma_i}^2}{\sum_{x \in X} w(x) \cdot {\mathrm{v}_x}^2}.$$

The variance of the error of prediction is estimated from the equation A3 increased by a factor $\frac{22}{21}$, considering the 22 ages (or equations) in the model and the degree of freedom lost after estimating $k_i^*$:

$$\widehat{\sigma_i}^2 = \sum_{x \in X} w(x) \cdot \epsilon_i(x)^2,$$

As a result, the variance of $k_i^*$ is given by equation A4:

$$\mathrm{Var}[k_i^*] = \frac{22}{21} \cdot \frac{\sum_{x \in X} w(x) \cdot \epsilon_i(x)^2}{\sum_{x \in X} w(x) \cdot \mathrm{v}_x{}^2}. \quad \text{(A5)}$$

After some additional steps, the variance of the estimator is redefined as a function of the residuals of the model when the pattern of the mortality is omitted $e_i(x)$ and the optimal value of $k_i$:

$$\mathrm{Var}[k_i^*] = \frac{22}{21} \cdot \left[\frac{\sum_{x \in X} w(x) \cdot e_i(x)^2}{\sum_{x \in X} w(x) \cdot \mathrm{v}_x{}^2} - k_i^{*2}\right]. \quad \text{(A6)}$$

Given the standard deviation of $k_i^*$, 95% confidence intervals where calculated assuming a normal distribution, to the form:

$$k_i^* \pm 1.96 \cdot \left[\mathrm{Var}[k_i^*]\right]^{1/2}.$$

## Appendix 2: General estimation of $h$ and $k$, using the method of Lagrange (nonlinear approach for numeric solutions)

The best linear estimator of $k_i$ (application of equation A4) assumes that the level of under-5 mortality $q_i(5y)$ and at least one other probability of dying $q_i(x)$ from 0 to $x$ are given. However, some applications require a general solution of $k_i$, when: *i)* $q_i(5y)$ is unknown; and/or *ii)* the log-quadratic model is used for matching/fitting specific functions that are not represented in the estimation, such as mortality rates, durations, and probabilities of dying that do not cumulate from zero (e.g., starting at some point after birth).

Inasmuch as some applications involve a transformation of the log-quadratic model, matching/fitting the log-quadratic model to some relevant data is a problem of optimization subject to nonlinear constraints. Hence, in the most general case, the relevant parameters $h_i$ and $k_i$ would result of solving a constrained optimization through numerical methods. From this perspective, the Lagrangian $\mathcal{L}$, represents a general problem of matching and optimization, using multipliers to add nonlinear constraints to the objective function, to the form:

$$\mathcal{L}(h_i, k_i, \lambda_i) = MSE(h_i, k_i) - \lambda_i \cdot \left[\ln[g(h_i, k_i)] - \ln[\bar{g}_i]\right], \qquad \text{(A7)}$$

where $MSE(h_i, k_i)$ is the mean squared error of a population $i$; $g(h_i, k_i)$ is the value to be matched, as a function of the parameters of the model modifying the level and pattern of the under-five mortality; $\bar{g}$ is the numerical value of the constraint that is specific to population $i$; and $\lambda_i$ is the Lagrange multiplier.

The general estimation of the model implies finding the values of $(h_i, k_i, \lambda_i)$ that will make the partial derivatives of equation A7 equal to zero. Using the Newton-Raphson approach, we multiply the gradient (vector of first derivatives) by the Moore-Penrose inverse of the Hessian (matrix of second derivatives) to adjust the values of an initial approximation. Assuming this approximation is relatively close to the true solution, the optimal values of $(h_i, k_i, \lambda_i)$ can be iteratively calculated by equation A8:

$$\begin{bmatrix} h_i^o \\ k_i^o \\ \lambda_i^o \end{bmatrix} = \begin{bmatrix} h_i \\ k_i \\ \lambda_i \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial h_i \partial h_i} & \frac{\partial^2 \mathcal{L}}{\partial h_i \partial k_i} & \frac{\partial^2 \mathcal{L}}{\partial h_i \partial \lambda_i} \\ \frac{\partial^2 \mathcal{L}}{\partial k_i \partial h_i} & \frac{\partial^2 \mathcal{L}}{\partial k_i \partial k_i} & \frac{\partial^2 \mathcal{L}}{\partial k_i \partial \lambda_i} \\ \frac{\partial^2 \mathcal{L}}{\partial \lambda_i \partial h_i} & \frac{\partial^2 \mathcal{L}}{\partial \lambda_i \partial k_i} & \frac{\partial^2 \mathcal{L}}{\partial \lambda_i \partial \lambda_i} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial h_i} \\ \frac{\partial \mathcal{L}}{\partial k_i} \\ \frac{\partial \mathcal{L}}{\partial \lambda_i} \end{bmatrix} , \qquad \text{(A8)}$$

given a set of partial derivatives are calculated numerically, using the following equations as examples:

$$\frac{\partial \mathcal{L}}{\partial h_i} \approx \frac{\mathcal{L}(h_i+\Delta, k_i, \lambda_i) - \mathcal{L}(h_i-\Delta, k_i, \lambda_i)}{2\cdot\Delta};$$

$$\frac{\partial \mathcal{L}}{\partial k_i} \approx \frac{\mathcal{L}(h_i, k_i+\Delta, \lambda_i) - \mathcal{L}(h_i, k_i-\Delta, \lambda_i)}{2\cdot\Delta};$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} \approx \frac{\mathcal{L}(h_i, k_i, \lambda_i+\Delta) - \mathcal{L}(h_i, k_i, \lambda_i-\Delta)}{2\cdot\Delta};$$

and

$$\frac{\partial^2 \mathcal{L}}{\partial h_i \partial k_i} \approx \frac{\mathcal{L}(h_i+\Delta, k_i+\Delta, \lambda_i) - \mathcal{L}(h_i-\Delta, k_i+\Delta, \lambda_i) - \mathcal{L}(h_i+\Delta, k_i-\Delta, \lambda_i) + \mathcal{L}(h_i-\Delta, k_i-\Delta, \lambda_i)}{4\cdot\Delta^2}.$$

Since matching two inputs is also a problem of optimization, equation A7 can be redefined to have two multipliers (one per matching constraint) and no minimization part involved $MSE = 0$, to the form:

$$\mathcal{L}(h_i, k_i, \lambda_{1,i}, \lambda_{2,i}) = -\lambda_{1,i} \cdot \left[\ln[g_1(h_i, k_i)] - \ln\left[\bar{g}_{1,i}\right]\right] - \lambda_{2,i} \cdot \left[\ln[g_2(h_i, k_i)] - \ln\left[\bar{g}_{2,i}\right]\right] \quad \text{(A9)}$$

Optimal solution of equation A9 is feasible using the same iterative procedure of equation A8. However, the gradient and the Hessian are augmented in one dimension in order to include the partial derivatives of the second multiplier.

Appendix Table A1: List of country-years included in the final Under-5 Mortality Database (U5MD) used for estimating the coefficients of the log-quadratic model

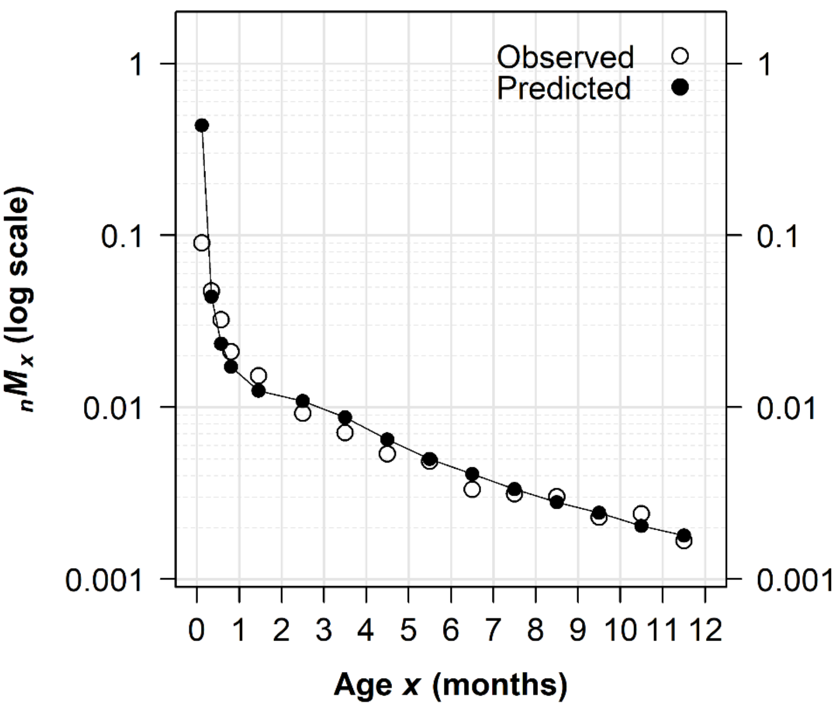
| Country | Years | n |
|---|---|---|
| Australia | 1935-1971, 1973-2014 | 79 |
| Austria | 1970-1994, 1996-2016 | 46 |
| Belgium | 1946-1954, 1956, 1961-1992, 2007-2010, 2013-2014 | 48 |
| Canada | 1929-1942, 1944-1975, 1977-1986, 1988-1990, 1992, 1995-1997, 1999-2006 | 71 |
| Chile | 1992-2004, 2007 | 14 |
| Denmark | 1928-1993, 1997, 2000-2015 | 83 |
| Finland | 1929-1940, 1946-1990, 1994, 1996-1998, 2000-2015 | 77 |
| France | 1953-1966, 1975-1992, 1996-1999, 2001-2015 | 51 |
| Germany | 1991-1994, 1996-1997, 2001-2007, 2010-2015 | 19 |
| West Germany | 1956-1960, 1970-1971, 1973-1977, 1979-1990 | 24 |
| Ireland | 1970-1988, 1990-1999, 2001-2006, 2008-2011 | 39 |
| Israel | 1983-1998, 2000-2016 | 33 |
| Italy | 1946-1955, 1957-1985, 1987-2013 | 66 |
| Japan | 1947-1950, 1954-1956, 1958-1959, 1963, 1970-1994, 1996-2000, 2002-2014 | 53 |
| Netherlands | 1970-1994, 1996, 1998, 2000-2001, 2004-2008 | 34 |
| New Zealand | 1970-1971, 1973-1975, 1977-2013 | 42 |
| Norway | 1935-1992, 1995-2001, 2003-2012 | 75 |
| Portugal | 1970-1993, 1996-1997, 2001-2015 | 41 |
| South Korea | 2004-2015 | 12 |
| Spain | 1976-1983, 1987-1991, 1995-1998, 2001-2013 | 30 |
| Sweden | 1934-2002, 2004-2012 | 78 |
| Switzerland | 1920-1930, 1970-1982, 1984-1994, 1996, 1998-2016 | 55 |
| United Kingdom | 1982-1991, 1993, 1996-2001, 2005-2012 | 25 |
| England and Wales | 1934-1985 | 52 |
| United States | 1933-1944, 1946-1993, 1995-1998, 2000-2003, 2008-2009, 2014-2015 | 72 |
| **Total** | | **1,219** |

Appendix Table A2: Mean Bias Error (MBE) of predicted q(x)'s using the log-quadratic model applied to the final U5MD with various combinations of outcomes and entry points for estimating k, both sexes combined

| | | MBE for the following outcomes: | | | |
|---|---|---|---|---|---|
| *Entry point(s)* | | *all q(x)* | *q(28d)* | *q(12m)* | *q(5y)* |
| *q(5y)* only, $k = 0$ | | 0.0000 | 0.0002 | 0.0001 | 0.0000 |
| | | *[-0.0026, 0.0027]* | *[-0.0161, 0.0168]* | *[-0.0052, 0.0053]* | - |
| *q(5y)* and | *q(7d)* | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | | *[-0.0019, 0.0019]* | *[-0.0052, 0.0051]* | *[-0.0044, 0.0043]* | - |
| | *q(28d)* | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | | *[-0.0017, 0.0017]* | - | *[-0.0043, 0.0042]* | - |
| | *q(3m)* | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | | *[-0.0013, 0.0013]* | *[-0.006, 0.006]* | *[-0.0035, 0.0034]* | - |
| | *q(6m)* | 0.0000 | -0.0001 | 0.0000 | 0.0000 |
| | | *[-0.0007, 0.0007]* | *[-0.0126, 0.0119]* | *[-0.002, 0.0021]* | - |
| | *q(12m)* | 0.0000 | -0.0002 | 0.0000 | 0.0000 |
| | | *[-0.001, 0.0009]* | *[-0.0192, 0.018]* | - | - |
| | all *q(x)* ✻ | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | | *[-0.0009, 0.0009]* | *[-0.0063, 0.0064]* | *[-0.0029, 0.0029]* | - |
| *q(28d, 5y)* only, $k = 0$ | | -0.0246 | -0.0266 | -0.0250 | -0.0241 |
| | | *[-0.0461, -0.0029]* | *[-0.0633, 0.0096]* | *[-0.0484, -0.0018]* | *[-0.0436, -0.005]* |

Reported values correspond to the mean of 10,000 random samples: 60% of the life tables were used for estimation and 40% for evaluation.
MBE calculated from the residuals of 487 life tables (40% of sample). 95% CIs reported.
✻ Using k = k* (Equation(4))

Appendix Figure A1: Observed and Predicted values of ${}_nM_x$ in Jordan, 2015, for both sexes combined




Observed = unadjusted VR-based ${}_nM_x$ values

Predicted = ${}_nM_x$ values predicted on the basis of the log-quadratic model with the observed VR-based value of $q$(28d,5y) and $k$ = 0 as inputs

## Supplementary Materials 1

# Description of the Under-5 Mortality Database (U5MD): data sources and methodology for calculating harmonized mortality rates by detailed age between 0 and 5 years

## Table of Contents

# 1. Introduction

The Under-5 Mortality Database (U5MD) used in the accompanying paper "Modeling Age Patterns of Under-5 Mortality: Results from a Log-Quadratic Model Applied to High-Quality Vital Registration Data" is a newly compiled database for under-5 mortality by detailed age drawn from high-quality Vital Registration (VR) data. This database contains annual distributions of under-5 deaths by sex and detailed age, representing 25 countries over a time window spreading from the second half of the nineteenth century to recent years. This information was used to calculate mortality rates and probabilities of dying by detailed age from birth to age 5. The U5MD is freely accessible at https://web.sas.upenn.edu/global-age-patterns-under-five-mortality/.

These Supplementary Materials 1 document: (1) the criteria for selecting countries; (2) the sources of information, including available variables and country-year exclusions due to insufficient information; (3) the procedure for harmonizing age intervals; and (4) the methods for estimating mortality indicators.

# 2. Selection of countries

We selected countries primarily on the basis of a data quality criterion, using the *Human Mortality Database* (HMD) as a reference. The HMD represents a benchmark for mortality estimates in terms of data quality. Thus, we decided to select countries only among those included in the HMD. The overlap between the HMD and the U5MD allowed us to use some of the relevant HMD information in the estimation procedure (see section 3.2).

One difference, however, is that unlike the HMD we did not include countries of the former Eastern Bloc. Numerous authors have pointed out the fact that the Soviet Union used criteria for defining and reporting live births and stillbirths that were different from international standards (Anderson et al. 1994; Anderson & Silver 1986; Davis & Feshbach 1980; Velkoff & Miller 1995). Indeed, the Soviet Union used a more restricted definition of a live birth in terms of weight and gestational age requirements. As a result, some live births that ended in death according to the international standard were counted as stillbirths according to the Soviet standard and thus excluded from mortality calculations. Another problem was the underregistration of infant deaths (among births meeting the Soviet definition of a live birth). This has led to significant amounts of understatement of the infant mortality rate in the Soviet Union but also in several European countries that were part of the Eastern Bloc and aligned on those definitions and practices (Gourbin & Masuy-Stroobant 1995). Studies have shown that underestimation of infant mortality continued after 1990, including after adopting international standards for the definition of a live birth (Aleshina & Redmond 2005; Guillot et al. 2013; Kingkade & Sawyer 2001). Despite recent improvements in the quality of infant mortality information in the region, we made the conservative decision to exclude all the countries of the former Eastern Bloc due to the critical importance of the mortality information at early ages in our study.

We also excluded Greece for similar reasons. According to the HMD report (Agorastakis et al. 2017), the country was affected by significant undercount of neonatal deaths at least until the 1980s. Finally, we removed Iceland and Luxembourg due to small population sizes leading to many zero cell counts in the narrow age intervals used in the database. After 2000, the number of deaths between 0 and 5 years tends to be less than 10 deaths per year in these two countries. In contrast, the average number of deaths tends to be higher than 100 per year in Scandinavian countries for the same period and higher than 1,000 in the other countries.

As a result of these country exclusions, the U5MD covers 25 countries, instead of 40 for the HMD. The list of countries included in the U5MD is presented in Table SM1-1.

# 3. Sources of information

The data come from two primary sources containing raw counts of under-5 deaths by detailed age: historical demographic/statistical yearbooks and a UN repository of vital statistics. In addition, we used the HMD as a secondary source to fill some gaps in the primary sources and supply information about exposures to the risk of death, i.e., the denominators of mortality rates.

## 3.1 Primary sources of information

First, we manually collected age distributions of deaths from archival sources such as national demographic or statistical yearbooks. Second, for the period 1970 onwards, we used age distributions of deaths from a data repository compiled by the United Nations Statistical Division. Table SM1-1 shows the country-years collected from each of these two primary sources. The original version of the U5MD contains 1,741 country-years. However, in order to build a fully harmonized database by sex and age, we excluded a number of country-years due to insufficient information.

The next subsections provide the main characteristics of these death distributions and the rationale for country-year exclusions.

### 3.1.1 Sex

Most of the death distributions that we collected in the original database were available by sex. We excluded 81 country-years that had information only for both sexes combined.

### 3.1.2 Time periods

The age distributions of deaths available in the primary sources were period-specific and mostly by calendar year (a few exceptions with time periods wider than one year are indicated in Table SM1-1). In the original database, the historical demographic/statistical yearbooks provided 1,004 age distributions of deaths from 1841 to 2001. The UN repository supplied 737 age distributions from 1970 to 2016.

As a result, the U5MD covers most of the Western experience of mortality transition, with a decline of under-5 mortality from around 400 to less than 5 deaths per 1,000 births.

### 3.1.3 Heterogeneity and harmonization of age intervals

#### *3.1.3.1 Heterogeneity of the length of age intervals*

The minimum age breakdown that we required was the distinction between neonatal (< 28 days or sometimes < 1 month) vs. post-neonatal deaths (rest of the first year). However, in most cases the available tabulations had greater age detail. The UN repository covered the first year of life with harmonized age intervals: by day for the first week, by week until the 28th day, and by month for the rest of the first year. By contrast, the age intervals available in the historical yearbooks were highly heterogeneous. The tables spanned the first five years of age unevenly: 615 age distributions of deaths covered the first year of life only, 170 the first two years, and 219 the first five years.

Age intervals were also reported through a multiplicity of formats. For example, in 1905, the yearbook of England and Wales reported neonatal deaths by week of age, and post-neonatal deaths by month of age. Starting in 1906, deaths occurring during the first day of life were separately shown. More details were introduced in 1931 when deaths occurring in the first week were reported by day of age (the information also included the number of deaths occurring within the first half hour of life). Since 1926, post-neonatal deaths have been grouped by trimester of age (with the exception of the period 1952-64 tabulated again by month). Figure SM1-1 shows the age distribution of deaths for the first year of life as it appears in *The Registrar General's Statistical Review of England and Wales for the year 1946*.

In Belgium and France, for many years in the 19th and 20th centuries, deaths were reported by 5-day age group for the first 10 or 15 days, and then by 5, 10, or 15-day age group for the rest of the first month. In both countries, the post-neonatal information was tabulated unevenly by month, trimester or semester. Figure SM1-2 shows the age distribution of deaths for the first five years of life for Belgium in 1924 provided in the *Annuaire Statistique de la Belgique*.

In some cases, few or no age details were reported for neonatal deaths. For example, from 1890 to 1920, Danish yearbooks provided the number of deaths in the first day of life as the only breakdown for the neonatal period (the yearbooks for the period 1896-1900 did not include breakdown at all). Information for the first week was added in 1921. Figure SM1-3 shows the age distribution of deaths provided for first three years of age in Denmark's *Statistisk Tablevierk* for that year.

There was no age breakdown within the first month of life in Belgium (1841-1861), Italy (1872-1889), New Zealand (1972), Portugal (1940-1954), and The Netherland (1850-1864).

For the post-neonatal period, the primary sources always include at minimum information by trimester or semester (with the exception of West Germany 1966-1988 with no age breakdown). After the first year, deaths were mostly reported by single year of age, with some exceptions for the second year: deaths were available by month in Australia (1921-1924) and by trimester in Belgium (1841-1861), The Netherlands (1850-1864), Norway (1876-1975), and Sweden (1891-1967).

Both the historical yearbooks and the UN repository have tabulations that include some death counts with unknown age. However, age was always identified at least by single year. In total, there are 203 country-years (including 158 from Norway) with some unknown-age deaths within single year age groups. The proportion of deaths with unknown age was small – less than 1% on average. We redistributed these deaths proportionally across the available detailed age intervals within single-year age groups.

#### *3.1.3.2 Heterogeneity of the age format*

In addition to variations in the length of age intervals, the age format also varied across sources. In the UN repository, death counts were uniformly formatted by day with months of 28, 60, 90, etc. days, that is with a year of 360 days. In the historical yearbooks, deaths were reported as integer by "hours'', "days'', "weeks", "months" or "years."

#### *3.1.3.3 Harmonizing age intervals*

In order to address the heterogeneity of the data, we harmonized age intervals in two ways. First, we recoded the original age formats: in order to estimate the precise exposure time to the risk of mortality, we assumed that the average duration of a year was 365.25 days (considering leap years of 366 days). Therefore, we set the average duration of a month to 30.4375 days (365.25/12). However, when the exact number of days of the first month was available (for example 28 days in Figure SM1-1 or 30 days in Figure SM1-2), we kept that exact duration and adjusted the duration of the second month accordingly.

Second, we harmonized the length of age intervals by week for the first 28 days, by month for the rest of the first year, by trimester for the second year, and by year for the three last years, using an interpolation method discussed in Section 4. At that stage, we excluded 78 country-years due to insufficient age details for performing this harmonization of age intervals.

### 3.1.4 Live births, stillbirths, and false stillbirths

The definition of a live birth has evolved throughout the 20th century. The early recommendation of the League of Nations in 1925 was the presence of breathing as vital sign to define a live birth. From 1950 onwards, the WHO replaced this recommendation by "any sign of life", making it more inclusive. This recommendation was progressively but unequally adopted by countries until today. For example, during the Soviet period, countries of the Eastern bloc often added the viability criteria of gestational duration and weight for identifying live births (Gourbin & Masuy-Stroobant 1995; Guillot et al. 2013). The definition of stillbirths varied accordingly.

In Belgium and France, a specific definition for stillbirths, adopted under France's First Empire, affected the measurement of mortality for certain time periods. In both countries, live births that had died before civil registration (legally within the first three days after delivery) were registered as stillbirths but tabulated separately from actual stillbirths. These "false stillbirths" were registered in Belgium (by sex) from 1879 to 1955 and from 1958 to 1960 (Glei, Devos, et al. 2017), and in France from 1899 to 1974 (by sex from 1953 to 1974) (Glei, Wilmoth, et al. 2017). Although we found a large number of death distributions by detailed age for these periods, we did not always find corresponding counts of false stillbirths. Therefore, we only included in the U5MD those years for which the false stillbirth counts were available by sex, that is from 1879 to 1954 in Belgium, and from 1953 to 1966 in France. We added these false stillbirths to the number of registered early-neonatal deaths (first week of age). These false stillbirths are thus taken into account in the estimation of mortality for those periods. 117 country-years were excluded from the database at that stage either because the false stillbirth count was not available at all, or because it was only available for both sexes combined.

## 3.2 Secondary source of information: The Human Mortality Database

We used the *Human Mortality Database* (HMD) as secondary source of information (Human Mortality Database 2018). The HMD is a public database that provides annual mortality and population data from birth to oldest ages by single year of age. The HMD provides data for 40 countries. The selection of countries was *"limited by design to populations where death registration and census data are virtually complete, since this type of information is required for the uniform method used to reconstruct historical data series. As a result, the countries and areas included here are relatively wealthy and for the most part highly industrialized".* The U5MD follows that criterion of virtual completeness by selecting, among the available data in the primary sources of information, only the country-years included in the HMD. Therefore, the U5MD can be seen as a complement to the HMD with greater age granularity at the early ages of life.

We used death counts from the HMD to fill potential missing information in the primary sources between exact ages 1 and 5. We also used the HMD exposures to the risk of dying for estimating mortality rates (see Section 5).

We extracted annual death counts and annual exposures to the risk of dying by single year of age. Note that, for both deaths and exposures, we combined the information for England and Wales, Scotland, and Northern Ireland to obtain the total counts for the UK. The total count for the UK is only available from 1982 onwards. Before 1982, the database provides the counts for England and Wales only. For consistency, we also aggregated the HMD data for the few periods that are longer than one year in the U5MD.

### 3.2.1 Death counts

We extracted death counts by sex for the first five years of life from the "input data files", i.e., the published raw deaths. These data were available in different Lexis areas (squares, triangles, and parallelograms) in the HMD primary sources. Therefore, we adopted the following rules to compute death counts by single year of age and one-year period (1x1 in the HMD terminology):

- When deaths were classified in lower and upper triangles, we added them up to obtain 1x1 Lexis squares.
- When deaths were classified by cohort in parallelograms centered on exact age ("VV" in HMD notation), we divided them into two triangles under assumption of uniform distribution. We then summed lower and upper triangles to obtain 1x1 Lexis squares.
- Some age intervals were not available by single year in the raw data. In these cases, we used the split age intervals of the HMD "complete data series" (adjusted death counts). More specifically, we used the relative age distributions of those adjusted deaths.
- We excluded data with LDB variable = 0. These correspond to data marked as "not used to create the Lexis database" in the HMD.
- When several death counts existed within the same 1x1 Lexis square, we summed them up to obtain the full 1x1 count.
- We ignored deaths with unknown age, which in the case of the HMD apply to the entire age range (0 to 100+).

### 3.2.2 Exposure to the risk of death

All exposure terms used in the U5MD were borrowed from the HMD "complete data series" with 1x1 Lexis squares. These exposures are expressed in person-years. In most cases, they correspond to mid-year population estimates derived from census enumerations assuming uniformity in the distribution of events (Wilmoth et al. 2017). However, when data on monthly births were available, the HMD team estimated person-years using this more detailed information (instead of making the assumption of uniformity in the distribution of births within a calendar year).

# 4. Harmonization of age intervals

All mortality estimates were computed for the same harmonized age intervals: weeks for the first 28 days, months for the rest of the first year, trimesters for the second year, and years for the last three years. This gives 22 age intervals with the following exact-age cut-off points: 0, 7, 14, 21, 28 days; 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 21 months; 2, 3, 4, 5 years.

Approaches for harmonizing age groups typically rely on interpolation methods. For example, to disaggregate death counts into single-year age groups, the HMD adopted a cubic spline interpolation method proposed by McNeil et al. (1977) applied to cumulative distributions of deaths. As noted in the HMD methods protocol (Wilmoth et al. 2017), the drawback of this approach is that the curve may sometimes decrease with age, generating negative counts of deaths. Such decreases with age are generally due to spurious oscillations created by the splines because of strong gradients in the data or non-equidistant points. To address this issue, which occur primarily at the oldest ages, the HMD protocol used specific constraints. However, the HMD approach does not completely eliminate negative deaths when applied to the U5MD. Therefore, we adopted an alternative interpolation method based on piecewise cubic interpolation or Hermite-type interpolation (Steffen 1990). This method guarantees a monotonic function in every case.

The method constructs a piecewise cubic interpolation function that passes through N given data points. It uses parabolas to determine the slope of the curve at an interval point *i* passing through points ($x_{i-1}$, $y_{i-1}$; $x_i$, $y_i$; $x_{i+1}$, $y_{i+1}$). Then a piecewise cubic function is constructed for each interval ($x_i$, $x_{i+1}$). With this approach, the slope of the curve has a continuous first-order derivative over the whole set of points. To ensure that the interpolation curves behave monotonically, the method verifies if the parabolas are monotonic. When it is not the case, the method takes as the slope the smallest of the two secants crossing either ($x_{i-1}$, $y_{i-1}$; $x_i$, $y_i$) or ($x_i$, $y_i$; $x_{i+1}$, $y_{i+1}$). For boundaries, the method uses the same approach but estimates the slope for the first (or last) point with parabolas fitted on the two next (or previous) points, and only using one secant.

We applied Steffen's method to cumulative distributions of deaths by sex. Harmonized deaths for both sexes combined were obtained by merging harmonized deaths by sex. In order to make the procedure robust, we applied the method only to the country-years that had at least a cut-off point between 7 days and 1 month (28 days or 30 days depending on the data format). We removed 78 country-years at that stage due to the absence of such cut-off point.

In total, the fully-harmonized database for which mortality rates were estimated thus contains 1,465 country years, i.e., 276 fewer than the original database (see Table SM1-1).

# 5. Mortality estimation

Using conventional notation, ${}_nM_x[t,t+1)$ is the mortality rate in the age interval $[x,x+n)$ and for the period $[t,t+1)$. $q(x)[t,t+1)$ is the probability of dying between age 0 and $x$ for the same period. Estimates are period-specific and mostly annual, i.e., they describe the mortality experience of a synthetic cohort for a given year.

We estimated age-specific mortality rates ${}_nM_x$ by dividing the number of deaths ${}_nD_x$ by the exposure to the risk of dying ${}_nE_x$ (expressed in person-years) for the period $[t,t+1)$, as shown in the following equation:

$${}_nM_x[t,t+1) = \frac{{}_nD_x\,[t,t+1)}{{}_nE_x\,[t,t+1)}$$

As explained in section 3.2, we extracted exposures to the risk of dying from the HMD by single year of age ${}_1E_{x'}$ where $x'$ is the lower bound of the age interval. In order to estimate the exposure for the detailed intervals of the U5MD (weeks, months, and trimesters), we assumed a uniform distribution of the exposure within the year interval. With this assumption, the exposure term is proportional to the length of each age interval $n$ (expressed in years):

$${}_nE_x[\mathrm{t},\mathrm{t}+1) = {}_1E_{x'}[\mathrm{t},\mathrm{t}+1) \cdot n$$

where $x' \leq x < x'+1$ and $n < 1$.

We then computed cumulative probabilities of dying $q(x)$ with the assumption that mortality rates were constant within each detailed age interval:

$$q(x) = 1 - e^{-\sum_{a=0}^{x-n} {}_nM_a \cdot n}$$

22 estimates of ${}_nM_x$ and $q(x)$ were thus calculated for each country-year, by sex and for both sexes combined: four estimates by week of age for the neonatal period; 11 estimates by month at post-neonatal ages; four estimates by trimester for the second year of life; and three estimates by single year of age for the remaining three years. This gives the following arrays:

For age-specific mortality rates, ${}_nM_x$:
${}_7M_{0(\text{d})}$, ${}_7M_{7(\text{d})}$, ${}_7M_{14(\text{d})}$, ${}_7M_{21(\text{d})}$,
${}_1M_{1(\text{m})}$, ${}_1M_{2(\text{m})}$, ${}_1M_{3(\text{m})}$, ${}_1M_{4(\text{m})}$, ${}_1M_{5(\text{m})}$, ${}_1M_{6(\text{m})}$, ${}_1M_{7(\text{m})}$, ${}_1M_{8(\text{m})}$, ${}_1M_{9(\text{m})}$, ${}_1M_{10(\text{m})}$, ${}_1M_{11(\text{m})}$,
${}_3M_{12(\text{m})}$, ${}_3M_{15(\text{m})}$, ${}_3M_{18(\text{m})}$, ${}_3M_{21(\text{m})}$,
${}_1M_{2(\text{y})}$, ${}_1M_{3(\text{y})}$, ${}_1M_{4(\text{y})}$

For cumulative probabilities of dying from birth to age $x$, $q(x)$:
$q$(7d), $q$(14d), $q$(21d), $q$(28d),
$q$(2m), $q$(3m), $q$(4m), $q$(5m), $q$(6m), $q$(7m), $q$(8m), $q$(9m), $q$(10m), $q$(11m), $q$(12m),
$q$(15m), $q$(18m), $q$(21m), $q$(24m)
$q$(3y), $q$(4y), $q$(5y)

## 6. References


Agorastakis, M., Jdanov, D., & Grigoriev, P. (2017). *About mortality data for Greece*. Retrieved from https://www.mortality.org/

Aleshina, N., & Redmond, G. (2005). How high is infant mortality in central and eastern Europe and the commonwealth of independent states? *Population Studies, 59*(1), 39-54. doi:10.1080/0032472052000332692

Anderson, B. A., Katus, K., & Silver, B. D. (1994). Developments and prospects for population statistics in countries of the former Soviet Union. *Population Index, 60*(1), 4-20. doi:10.2307/3645322

Anderson, B. A., & Silver, B. D. (1986). Infant mortality in the Soviet Union: Regional differences and measurement issues. *Population and Development Review, 12*(4), 705-738. doi:10.2307/1973432

Davis, C., & Feshbach, M. (1980). *Rising infant mortality in the USSR in the 1970's* (Vol. No. 74). Washington, DC: US Bureau of the Census.

Glei, D. A., Devos, I., & Poulain, M. (2017). *About mortality data for Belgium*. Retrieved from https://www.mortality.org/

Glei, D. A., Wilmoth, J., Vallin, M., Vallin, J., & Meslé, F. (2017). *About mortality data for France, total population*. Retrieved from https://www.mortality.org/

Gourbin, C., & Masuy-Stroobant, G. (1995). Registration of vital data: are live births and stillbirths comparable all over Europe? *Bulletin of the World Health Organization, 73*(4), 449-460.

Guillot, M., Lim, S.-j., Torgasheva, L., & Denisenko, M. (2013). Infant mortality in Kyrgyzstan before and after the break-up of the Soviet Union. *Population Studies, 67*(3), 335-352. doi:10.1080/00324728.2013.835859

Human Mortality Database. (2018). University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany), available at www.mortality.org (data downloaded on November 20, 2018).

Kingkade, W. W., & Sawyer, C. C. (2001). *Infant mortality in Eastern Europe and the former Soviet Union before and after the breakup*. Paper presented at the 2001 International Population Conference of the International Union for the Scientific Study of Population, Salvador de Bahia (Brazil).

McNeil, D. R., Trussell, T. J., & Turner, J. C. (1977). Spline interpolation of demographic data. *Demography, 14*(2), 245-252. doi:10.2307/2060581

Steffen, M. (1990). A simple method for monotonic interpolation in one dimension. *Astronomy and Astrophysics, 239*, 443-450.

Velkoff, V. A., & Miller, J. E. (1995). Trends and differentials in infant mortality in the Soviet Union, 1970-90: How much is due to misreporting? *Population Studies, 49*(2), 241-258. doi:10.1080/0032472031000148496

Wilmoth, J. R., Andreev, K., Jdanov, D., Glei, D. A., & Riffe, T. (2017). *Methods protocol for the human mortality database*. Retrieved from https://www.mortality.org/Public/Docs/MethodsProtocol.pdf

Table SM1-1. Country-years included in the original and the harmonized versions of the Under-5 Mortality Database (U5MD)

| Country | Original VR database | | | | Excluded country-years due to insufficient information on: | | | Fully harmo-nized database |
|---|---|---|---|---|---|---|---|---|
| | Years | Sources: Statistical Yearbooks | UN Database | Total | Sex | False still-births | Age | Total |
| Australia | 1921-71, 1973-2014 | 51 | 42 | 93 | | | | 93 |
| Austria | 1970-94, 1996-2016 | - | 46 | 46 | | | | 46 |
| Belgium | 1841-60, 1861-70*, 1878-84, 1886-1913, 1919-56, 1958-92, 2007-10, 2013-14 | 113 | 22 | 135 | 33 | 26 | (21**) | 76 |
| Canada | 1929-42, 1944-75, 1977-86, 1988-90, 1992, 1995-97, 1999-2006 | 41 | 30 | 71 | | | | 71 |
| Chile | 1992-2005, 2007 | - | 14 | 14 | | | | 14 |
| Denmark | 1890-94, 1896-1993, 1997, 2000-2015 | 79 | 41 | 120 | | | 30 | 90 |
| Finland | 1878-1920, 1921-25*, 1926-40, 1946-90, 1994, 1996-98, 2000-15 | 91 | 33 | 124 | | | | 124 |
| France | 1855-68, 1877-1889, 1891-1947, 1950-66, 1970-72, 1974-92, 1996-99, 2001-15 | 101 | 41 | 142 | | 91 | | 51 |
| Germany | 1991-94, 1996-97, 2001-07, 2010-15 | - | 19 | 19 | | | | 19 |
| West Germany | 1956-60, 1970-71, 1973-77, 1979-90 | 5 | 19 | 24 | | | | 24 |
| Ireland | 1970-88, 1990-99, 2001-06, 2008-11 | - | 39 | 39 | | | | 39 |
| Israel | 1983-98, 2000-16 | - | 33 | 33 | | | | 33 |
| Italy | 1872-89, 1926-33, 1939-55, 1957-85, 1987-2013 | 85 | 14 | 99 | | | 18 | 81 |
| Japan | 1947-50, 1954-56, 1958-59, 1963, 1970-94, 1996-2000, 2002-14 | 10 | 43 | 53 | | | | 53 |
| Netherlands | 1850-64, 1970-94, 1996, 1998, 2000-01, 2004-08 | 15 | 34 | 49 | | | 15 | 34 |
| New Zealand | 1970-75, 1977-2013 | - | 43 | 43 | | | 1 | 42 |
| Norway | 1876-1900, 1901-05*, 1906-26, 1927-30*, 1931-92, 1995-2001, 2003-12 | 93 | 34 | 127 | | | | 127 |
| Portugal | 1940, 1942-59, 1962, 1970-93, 1996-97, 2001-15 | 20 | 41 | 61 | | | 14 | 47 |
| South Korea | 2004-15 | - | 12 | 12 | | | | 12 |
| Spain | 1976-83, 1987-91, 1995-98, 2001-13 | - | 30 | 30 | | | | 30 |
| Sweden | 1891-02, 2004-12 | 111 | 10 | 121 | | | | 121 |
| Switzerland | 1877-79, 1882-83, 1911-82, 1984-94, 1996, 1998-2016 | 64 | 44 | 108 | 48 | | | 60 |
| United Kingdom | 1982-91, 1993, 1996-2001, 2005-12 | - | 25 | 25 | | | | 25 |
| England and Wales | 1905-1985 | 65 | 16 | 81 | | | | 81 |
| United States | 1933-44, 1946-93, 1995-98, 2000-03, 2008-09, 2014-15 | 60 | 12 | 72 | | | | 72 |
| **Total** | | **1,004** | **737** | **1,741** | **81** | **117** | **78** | **1,465** |

* Deaths aggregated over the indicated range of years

** Country-years first excluded due to lack of data by sex

Figure SM1-1. Age distribution of infant deaths in The Registrar General's Statistical Review of England and Wales for the year 1946

**TABLE 13.—Deaths at Various Periods in the First Year of Life, 1946, and the Four Quarters thereof. {England and Wales, Geographical Regions, Aggregates of County Boroughs, Other Urban Districts and Rural Districts.**

*See Explanatory notes on page* v.

| | | | Total under 1 year | Under 30 minutes | 30 minutes and under 1 day | Total under 1 day | Days 1 | Days 2 | Days 3 | Days 4 | Days 5 | Days 6 | 1 day and under 1 week | Weeks 0 | Weeks 1 | Weeks 2 | Weeks 3 | Months: Total under 4 weeks | Months: 4 weeks to 3 mths. | Months 3-6 | Months 6-9 | Months 9-12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| England and Wales. | All Infants ... ... ... | M | 19458 | 518 | 3607 | 4125 | 1444 | 1109 | 746 | 432 | 312 | 324 | 4367 | 8492 | 1483 | 903 | 719 | 11597 | 3286 | 2555 | 1314 | 706 |
| | | F | 14083 | 453 | 2495 | 2948 | 985 | 681 | 484 | 336 | 292 | 236 | 3014 | 5962 | 1086 | 683 | 529 | 8260 | 2285 | 1931 | 1033 | 574 |
| | | P | 33541 | 971 | 6102 | 7073 | 2429 | 1790 | 1230 | 768 | 604 | 560 | 7381 | 14454 | 2569 | 1586 | 1248 | 19857 | 5571 | 4486 | 2347 | 1280 |
| | Legitimate ... ... ... | M | 17526 | 407 | 3346 | 3753 | 1337 | 1014 | 690 | 398 | 286 | 303 | 4028 | 7781 | 1344 | 818 | 639 | 10582 | 2865 | 2259 | 1175 | 645 |
| | | F | 12677 | 325 | 2276 | 2601 | 903 | 631 | 453 | 307 | 273 | 221 | 2788 | 5389 | 982 | 626 | 474 | 7471 | 2011 | 1739 | 930 | 526 |
| | | P | 30203 | 732 | 5622 | 6354 | 2240 | 1645 | 1143 | 705 | 559 | 524 | 6816 | 13170 | 2326 | 1444 | 1113 | 18053 | 4876 | 3998 | 2105 | 1171 |
| | Illegitimate ... ... ... | M | 1932 | 111 | 261 | 372 | 107 | 95 | 56 | 34 | 26 | 21 | 339 | 711 | 139 | 85 | 80 | 1015 | 421 | 296 | 139 | 61 |
| | | F | 1406 | 128 | 219 | 347 | 82 | 50 | 31 | 29 | 19 | 15 | 226 | 573 | 104 | 57 | 55 | 789 | 274 | 192 | 103 | .48 |
| | | P | 3338 | 239 | 480 | 719 | 189 | 145 | 87 | 63 | 45 | 36 | 565 | 1284 | 243 | 142 | 135 | 1804 | 695 | 488 | 242 | 109 |

Figure SM1-2. Age distribution of under-5 deaths in Belgium's *Annuaire Statistique de la Belgique* for the year 1924

48

DÉCÈS PAR AGE (*).

(MORT-NÉS ET AUTRES ENFANTS PRÉSENTÉS SANS VIE NON COMPRIS.)

| AGE au MOMENT DU DÉCÈS. | PROPORTION PAR MILLE DE CHAQUE AGE. 1921 | 1922 | 1923 | 1924 CHIFFRES ABSOLUS. Sexe masculin. | | Sexe féminin. | | TOTAL. | | Prop. par 1,000 de chaque âge. Sexe masculin. | | Sexe féminin. | | TOTAL | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Moins de 5 jours | 17.41 | 16.16 | 15.32 | 7,650 | 871 | 5,968 | 676 | 13,618 | 1,547 | 148.40 | 16.90 | 124.08 | 14.06 | 136.65 | 15.52 |
| 5 à moins de 10 jours | 9.55 | 7.72 | 7.65 | | 425 | | 308 | | 733 | | 8.24 | | 6.40 | | 7.35 |
| 10 — 20 — | 16.73 | 12.03 | 12.17 | | 610 | | 476 | | 1,086 | | 11.83 | | 9.90 | | 10.90 |
| 20 — 30 — | 12.20 | 8.40 | 8.87 | | 403 | | 348 | | 751 | | 7.82 | | 7.23 | | 7.54 |
| 1 — 2 mois | 20.28 | 15.84 | 16.38 | | 802 | | 614 | | 1,416 | | 15.56 | | 12.77 | | 14.21 |
| 2 — 3 — | 20.96 | 16.13 | 15.43 | | 782 | | 609 | | 1,391 | | 15.17 | | 12.66 | | 13.96 |
| 3 — 6 — | 41.51 | 36.08 | 31.18 | | 1,693 | | 1,285 | | 2,978 | | 32.84 | | 26.72 | | 29.88 |
| 6 — 12 — | 44.36 | 41.83 | 37.71 | | 2,064 | | 1,652 | | 3,716 | | 40.04 | | 34.34 | | 37.29 |
| 1 — 2 ans | 29.00 | 30.55 | 26.87 | 2,323 | 1,281 | 1,971 | 1,072 | 4,294 | 2,353 | 45.06 | 24.85 | 40.97 | 22.28 | 43.09 | 23.61 |
| 2 — 3 — | 8.33 | 12.17 | 10.77 | | 504 | | 409 | | 913 | | 9.78 | | 8.50 | | 9.16 |
| 3 — 4 — | 4.59 | 4.78 | 6.58 | | 326 | | 260 | | 586 | | 6.32 | | 5.41 | | 5.88 |
| 4 — 5 — | 4.30 | 2.99 | 3.29 | | 212 | | 230 | | 442 | | 4.11 | | 4.78 | | 4.44 |

Figure SM1-3. Age distribution of deaths under age 3 in Denmark's *Statistisk Tablevierk* for the year 1921

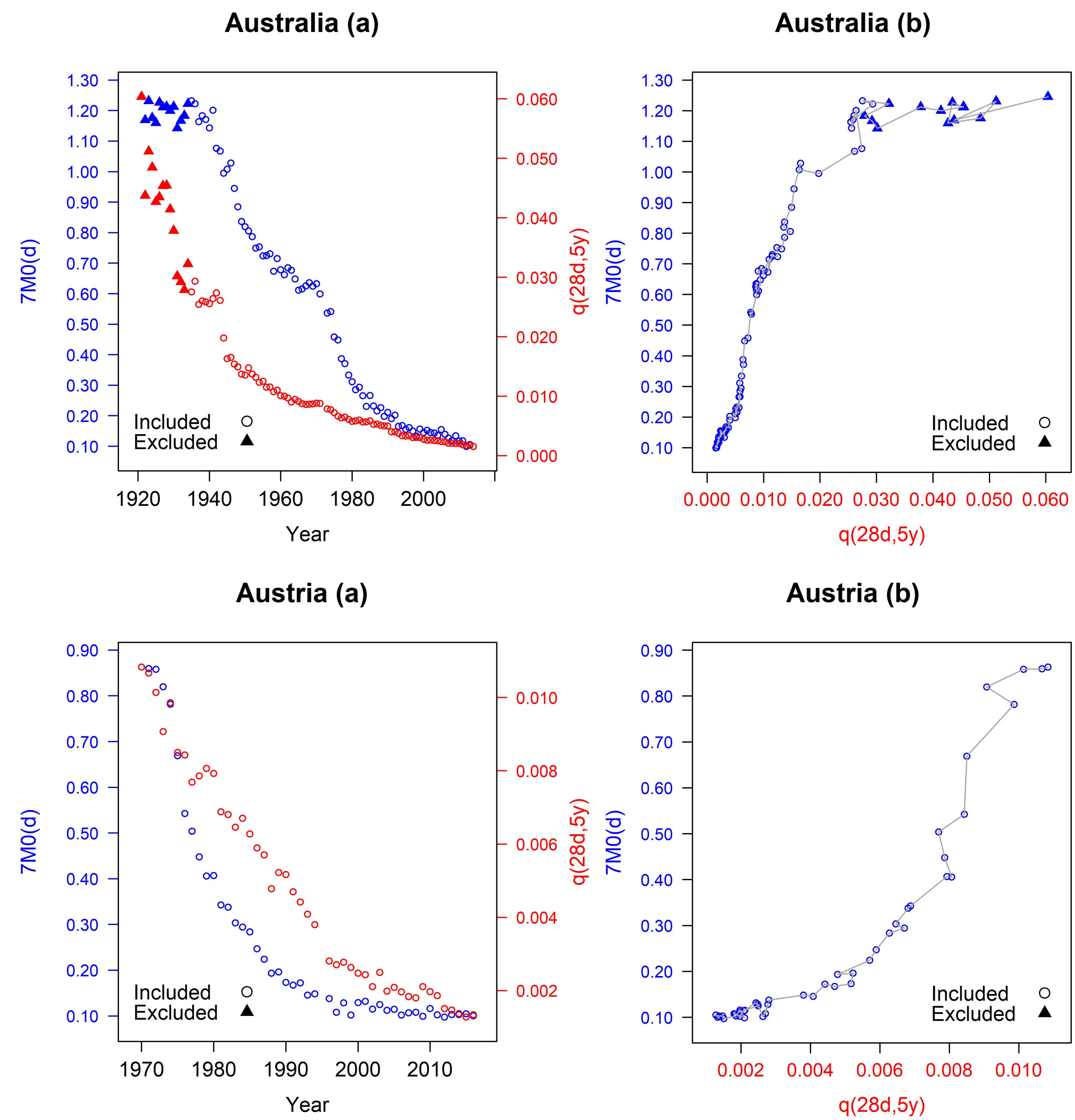
**Tabel IV E. Børnedød**

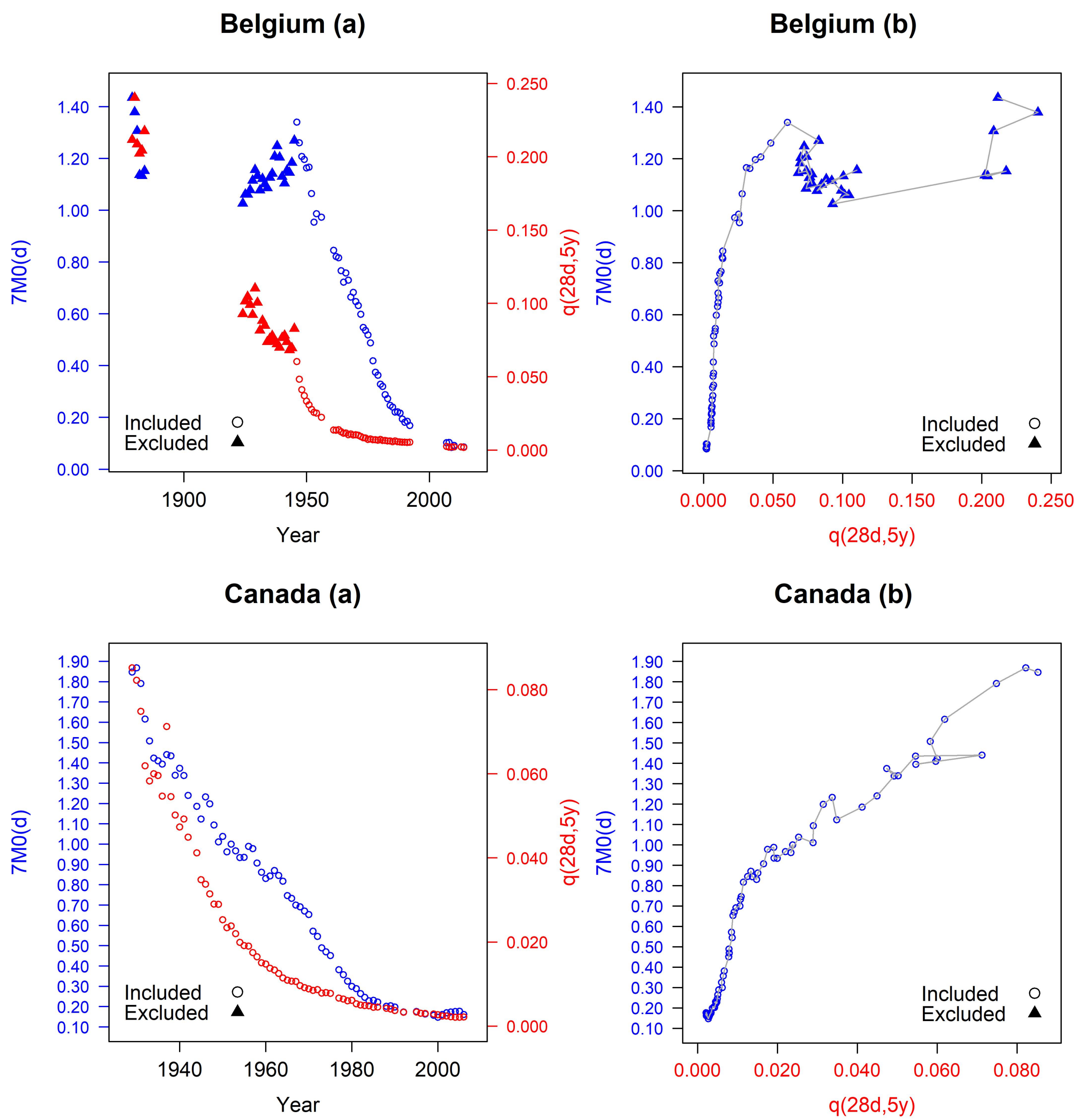
*Tableau IV E. Mortalit*

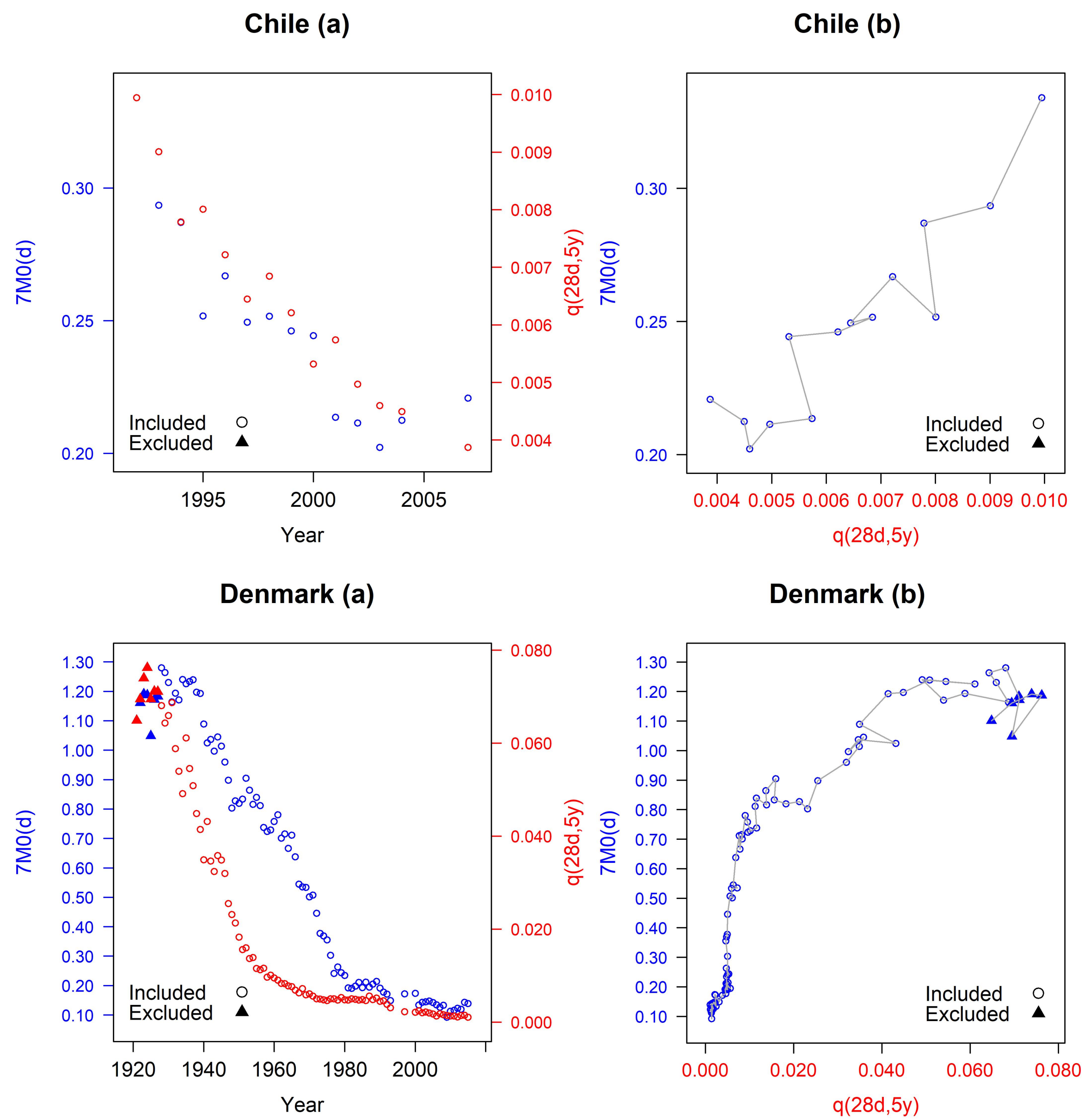
| Aar *ans* | Drenge *garcons* | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Under 24 Timer *au-des-sous de 24 heures* | 1—6 Dage *1—6 jours* | 7 Dage til under 1 Md. *7 jours—1 mois* | 1 Md. *1 mois* | 2 Md. *2 mois* | 3 Md. *3 mois* | 4 Md. *4 mois* | 5 Md. *5 mois* | 6 Md. *6 mois* | 7 Md. *7 mois* | 8 Md. *8 mois* | 9 Md. *9 mois* | 10 Md. *10 mois* | 11 Md. *11 mois* | Tils. under 1 Aar *total moins d'un an* | 1 Aar *1 an* | 2 Aar *2 ans* |
| 1921 | | | | | | | | | | | | | | | | | |
| Januar | 41 | 35 | 38 | 38 | 29 | 27 | 21 | 20 | 18 | 17 | 16 | 11 | 4 | 6 | 321 | 34 | 17 |
| Februar | 25 | 33 | 50 | 34 | 37 | 26 | 28 | 23 | 25 | 27 | 18 | 15 | 21 | 8 | 370 | 41 | 25 |
| Marts | 35 | 45 | 54 | 38 | 42 | 32 | 24 | 25 | 24 | 21 | 18 | 12 | 12 | 11 | 393 | 58 | 17 |
| April | 43 | 51 | 50 | 44 | 38 | 30 | 25 | 21 | 16 | 19 | 23 | 25 | 17 | 12 | 414 | 39 | 16 |
| Maj | 46 | 47 | 40 | 40 | 23 | 30 | 18 | 12 | 10 | 9 | 7 | 6 | 7 | 13 | 308 | 50 | 13 |
| Juni | 30 | 41 | 31 | 32 | 28 | 20 | 7 | 9 | 9 | 4 | 4 | 2 | 9 | 8 | 234 | 21 | 12 |
| Juli | 38 | 31 | 42 | 28 | 29 | 19 | 12 | 8 | 10 | 5 | 3 | 8 | 5 | 4 | 242 | 23 | 10 |
| August | 49 | 42 | 29 | 22 | 24 | 18 | 12 | 7 | 9 | 5 | 3 | 5 | 3 | 3 | 231 | 23 | 3 |
| September | 35 | 31 | 32 | 32 | 18 | 16 | 14 | 4 | 6 | 6 | 5 | 3 | 1 | 2 | 205 | 20 | 17 |
| Oktober | 38 | 43 | 35 | 27 | 17 | 19 | 12 | 10 | 8 | 2 | 7 | 6 | 1 | 4 | 229 | 17 | 7 |
| November | 43 | 31 | 30 | 21 | 19 | 11 | 11 | 14 | 7 | 10 | 3 | 3 | 3 | 2 | 208 | 16 | 10 |
| December | 50 | 42 | 36 | 37 | 31 | 25 | 22 | 17 | 16 | 14 | 7 | 10 | 5 | 4 | 316 | 16 | 12 |
| Tilsammen *total* | 473 | 472 | 467 | 393 | 335 | 273 | 206 | 170 | 158 | 139 | 114 | 106 | 88 | 77 | 3 471 | 358 | 159 |

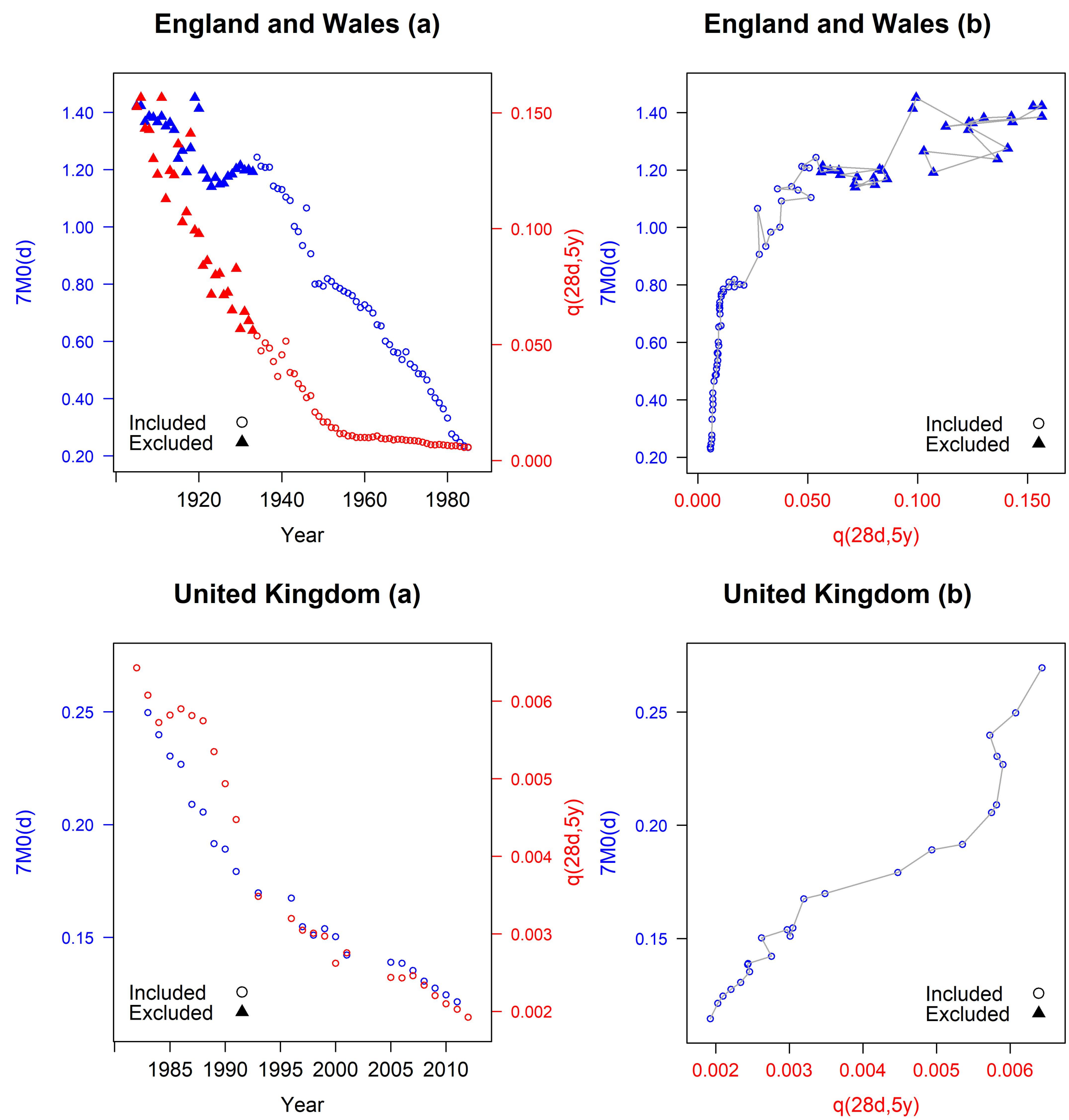
# Supplementary Materials 2: Additional Figures

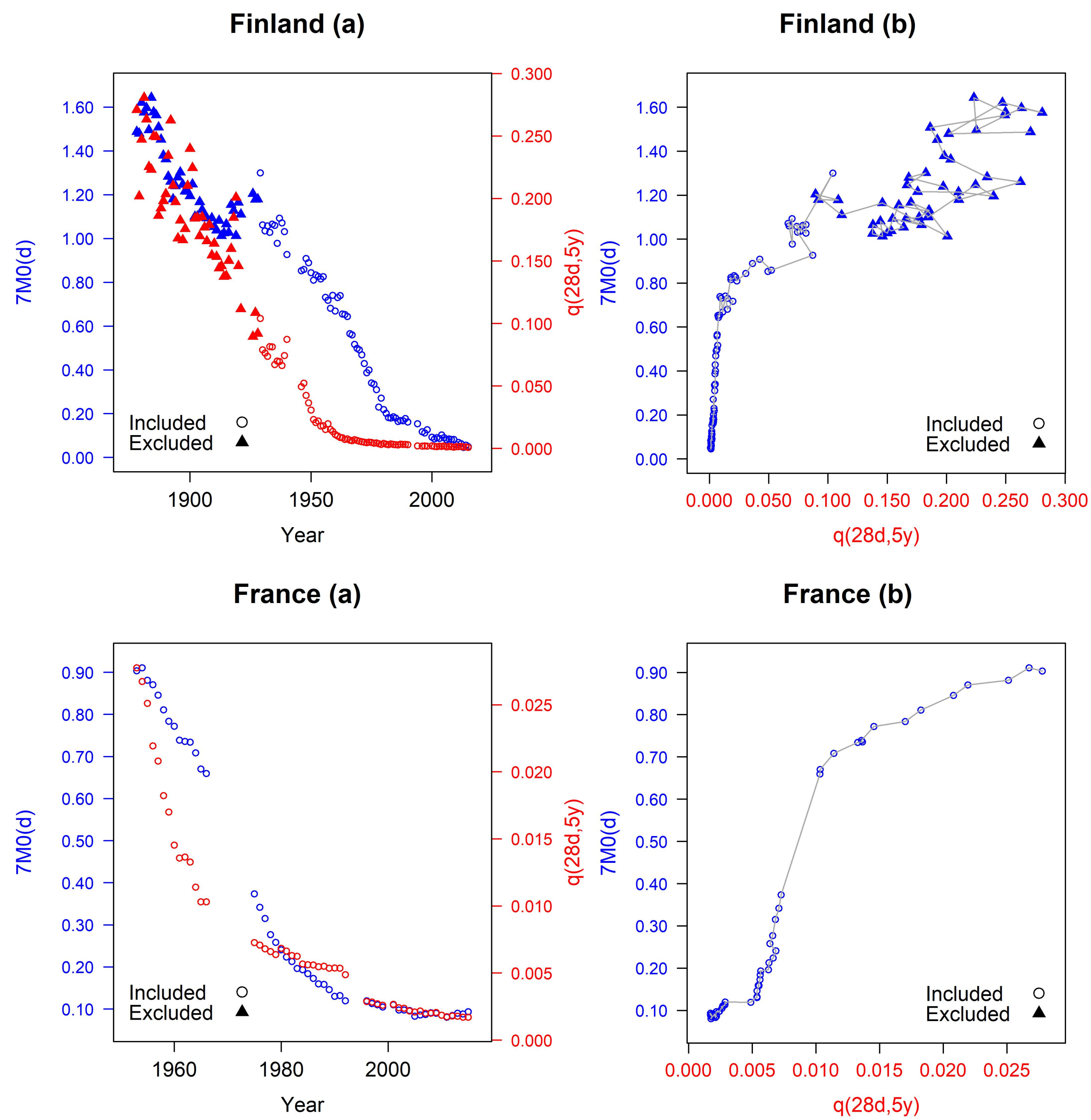
Figure SM2-1: Included and excluded country-years in the final U5MD according to (a) the time trends of 7M0(d) and q(28d,5y), and (b) the relationship between 7M0(d) and q(28d,5y), both sexes combined, by country

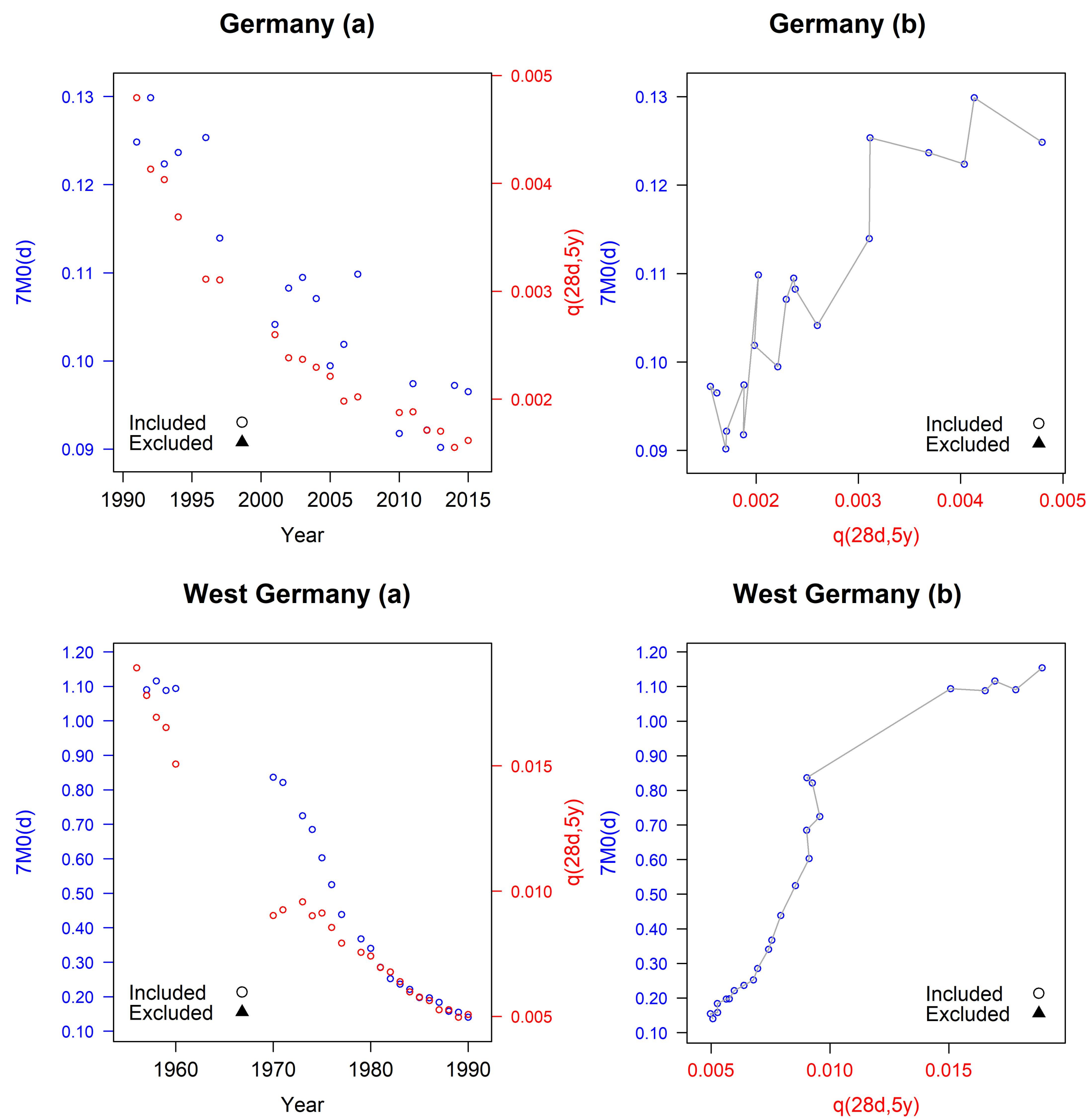


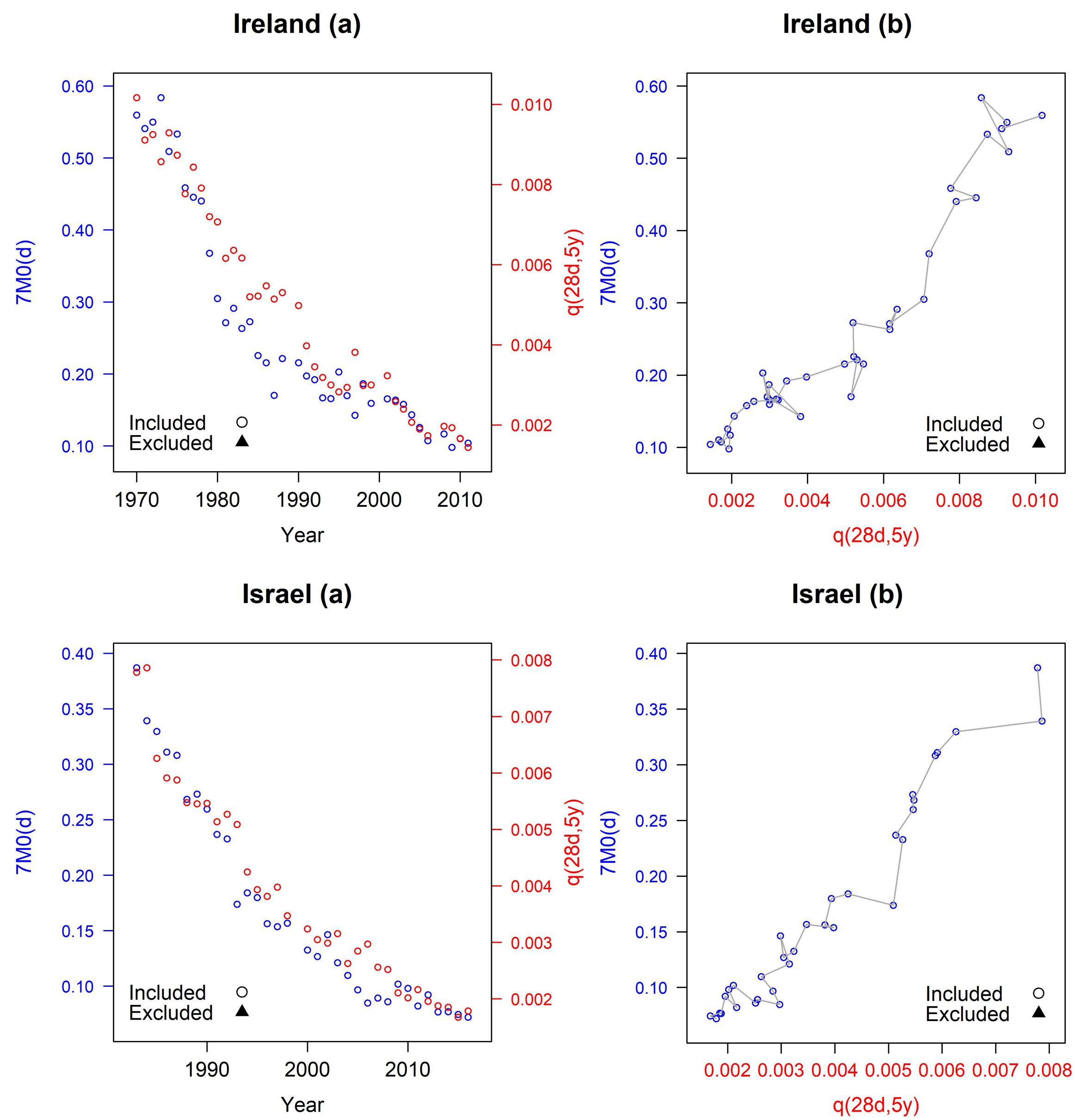

Belgium (a)
Belgium (b)
Canada (a)
Canada (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0.250
0.200
0.150
0.100
0.050
0.000
1900
1950
2000
1.90
1.80
1.70
1.60
1.50
1.40
1.30
1.20
1.10
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.080
0.060
0.040
0.020
0.000
1940
1960
1980
2000

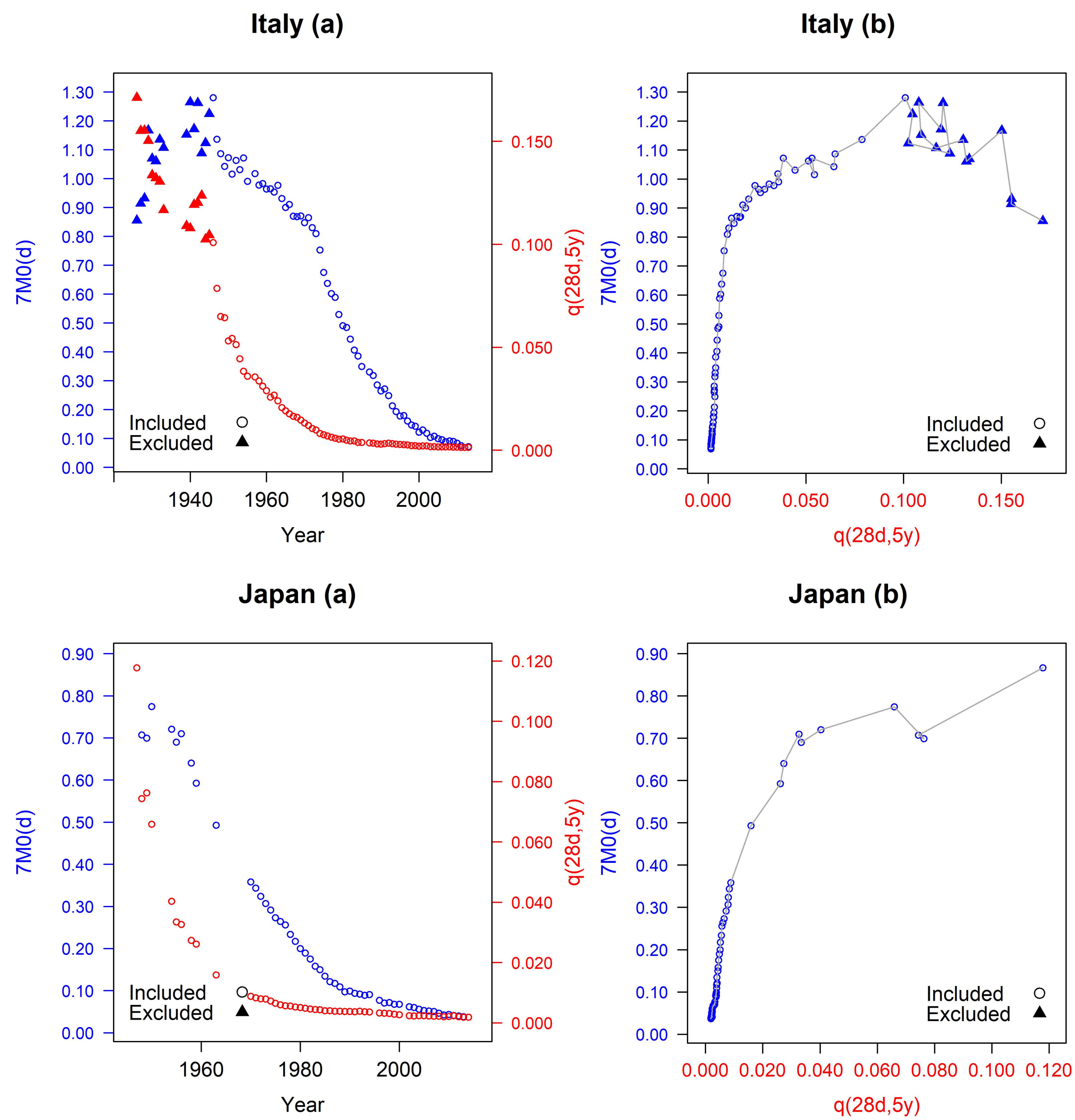

Chile (a)
Chile (b)
Denmark (a)
Denmark (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
0.20
0.25
0.30
0.004
0.005
0.006
0.007
0.008
0.009
0.010
1995
2000
2005
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
1.30
0.000
0.020
0.040
0.060
0.080
1920
1940
1960
1980
2000


England and Wales (a)
England and Wales (b)
United Kingdom (a)
United Kingdom (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.150
0.100
0.050
0.000
1920
1940
1960
1980
0.25
0.20
0.15
0.006
0.005
0.004
0.003
0.002
1985
1990
1995
2000
2005
2010

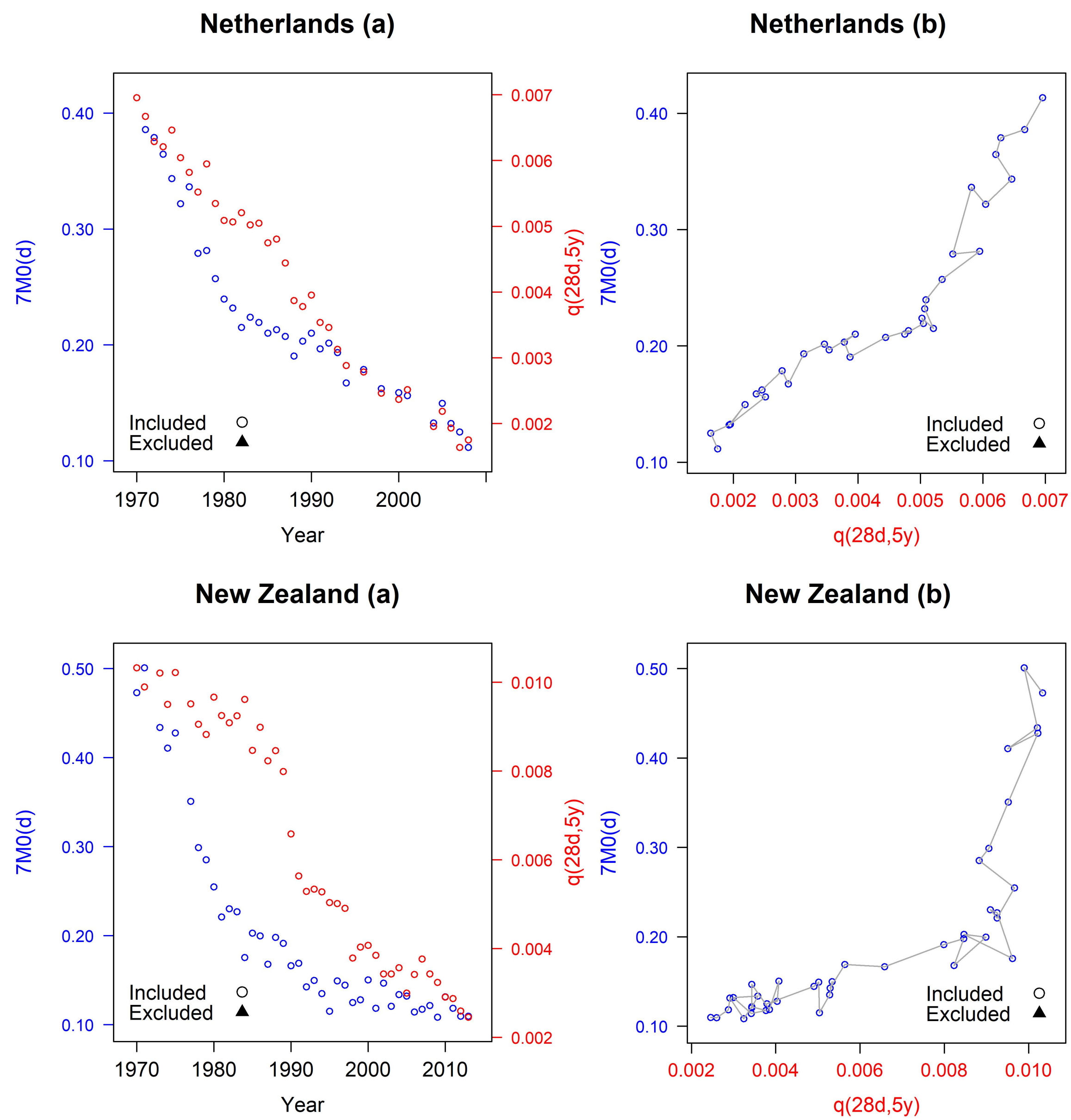

Finland (a)
Finland (b)
France (a)
France (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded


Germany (a)
Germany (b)
West Germany (a)
West Germany (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
0.09
0.10
0.11
0.12
0.13
0.002
0.003
0.004
0.005
1990
1995
2000
2005
2010
2015
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
0.005
0.010
0.015
1960
1970
1980
1990

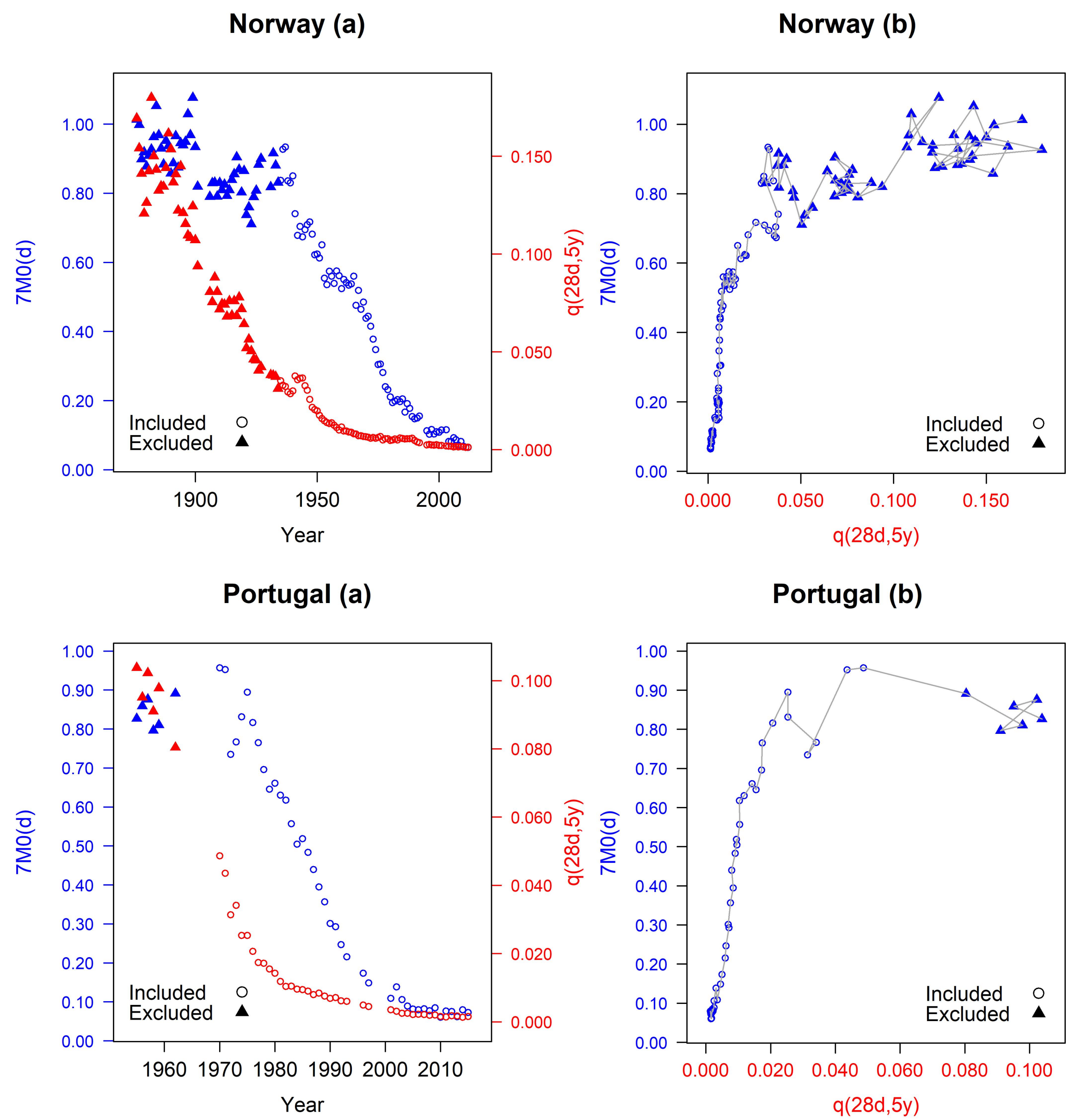

Ireland (a)
Ireland (b)
Israel (a)
Israel (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
0.60
0.50
0.40
0.30
0.20
0.10
0.010
0.008
0.006
0.004
0.002
1970
1980
1990
2000
2010
0.40
0.35
0.25
0.15
0.007
0.005
0.003

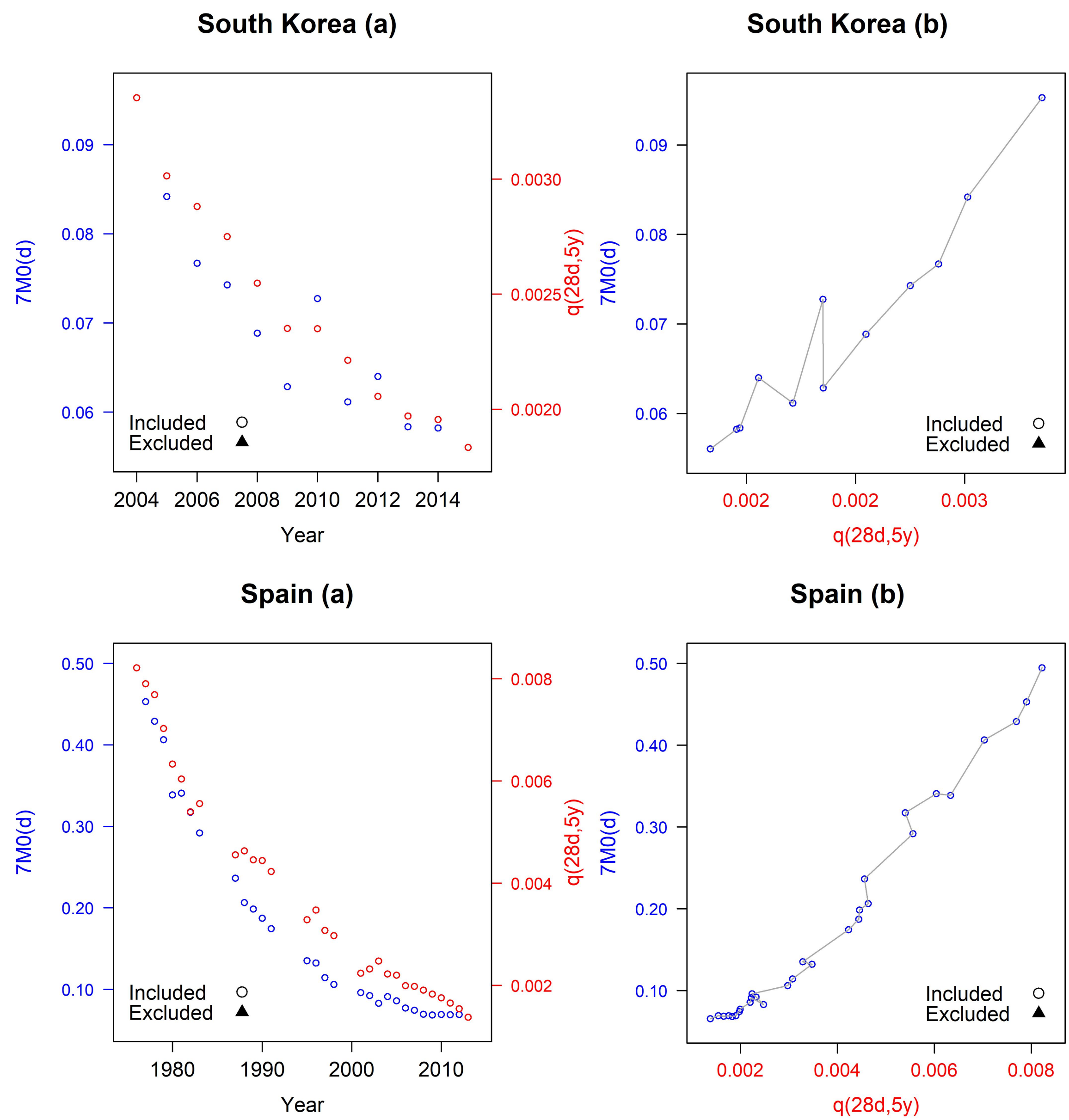


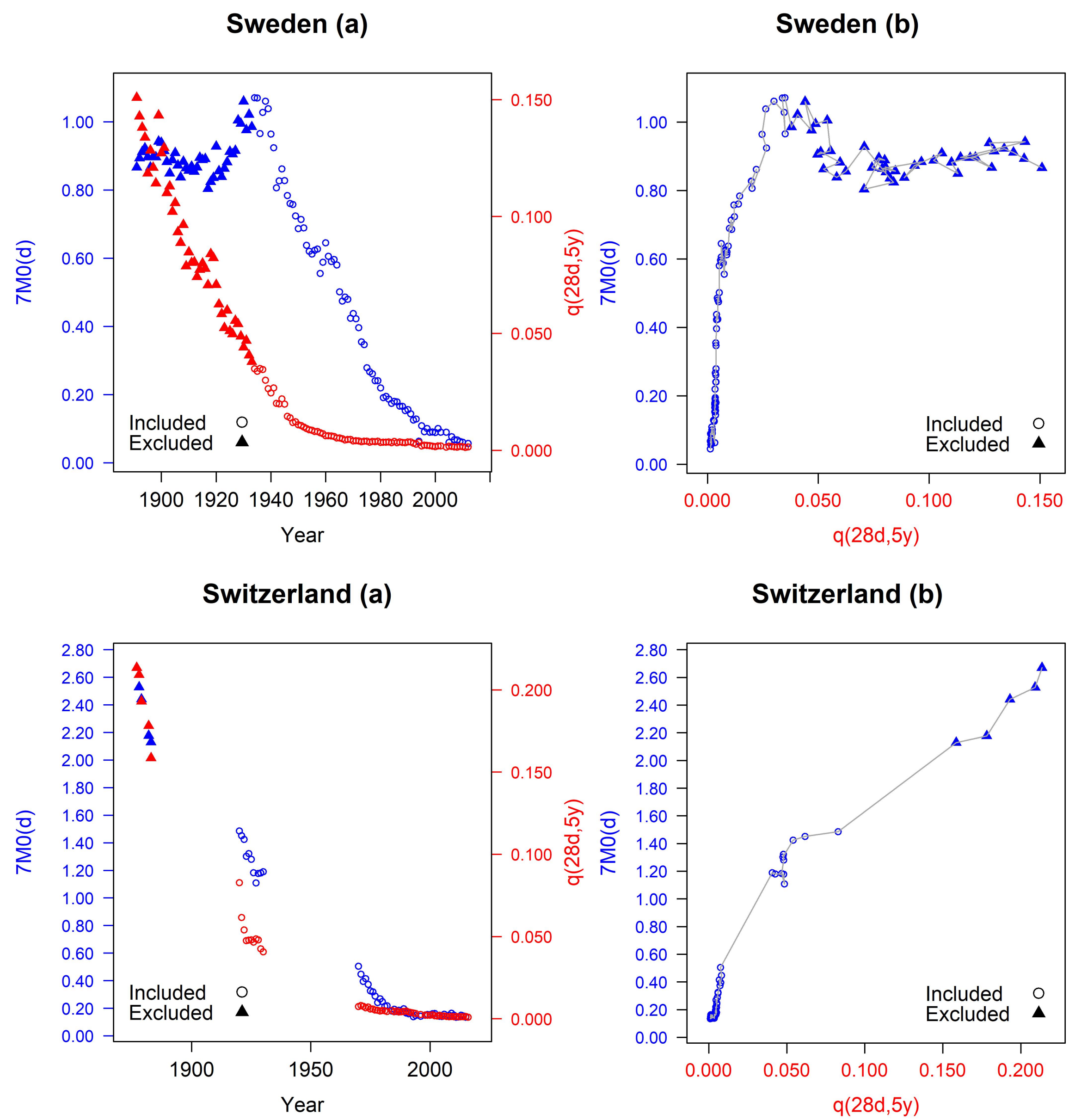
Italy (a)
Italy (b)
Japan (a)
Japan (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
1940
1960
1980
2000
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10
1.20
1.30
0.000
0.050
0.100
0.150
0.020
0.040
0.060
0.080
0.120



Netherlands (a)
Netherlands (b)
New Zealand (a)
New Zealand (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
0.40
0.30
0.20
0.10
0.007
0.006
0.005
0.004
0.003
0.002
1970
1980
1990
2000
0.50
0.010
0.008
0.006
0.004
0.002
2010


Norway (a)
Norway (b)
Portugal (a)
Portugal (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded


South Korea (a)
South Korea (b)
Spain (a)
Spain (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded
0.09
0.08
0.07
0.06
0.0030
0.0025
0.0020
2004
2006
2008
2010
2012
2014
0.002
0.003
0.50
0.40
0.30
0.20
0.10
0.008
0.006
0.004
1980
1990
2000
2010


Sweden (a)
Sweden (b)
Switzerland (a)
Switzerland (b)
7M0(d)
q(28d,5y)
Year
Included
Excluded

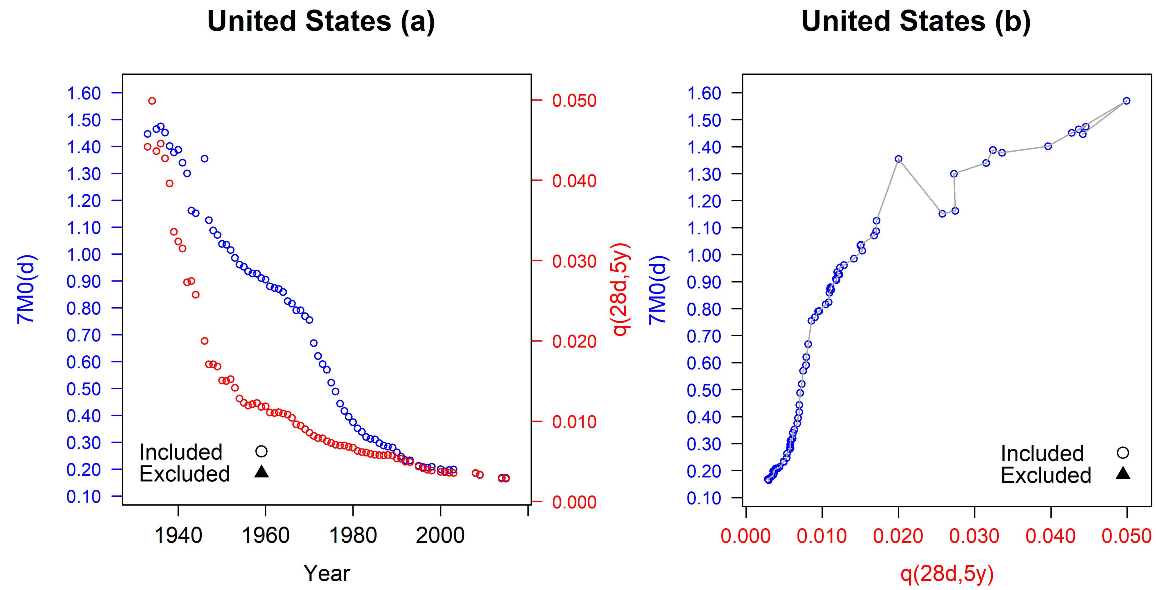



Figure SM2-2: Included and excluded country-years in the final U5MD according to (a) the time trends of 7M0(d) and q(28d,5y), and (b) the relationship between 7M0(d) and q(28d,5y), both sexes combined, all 25 countries

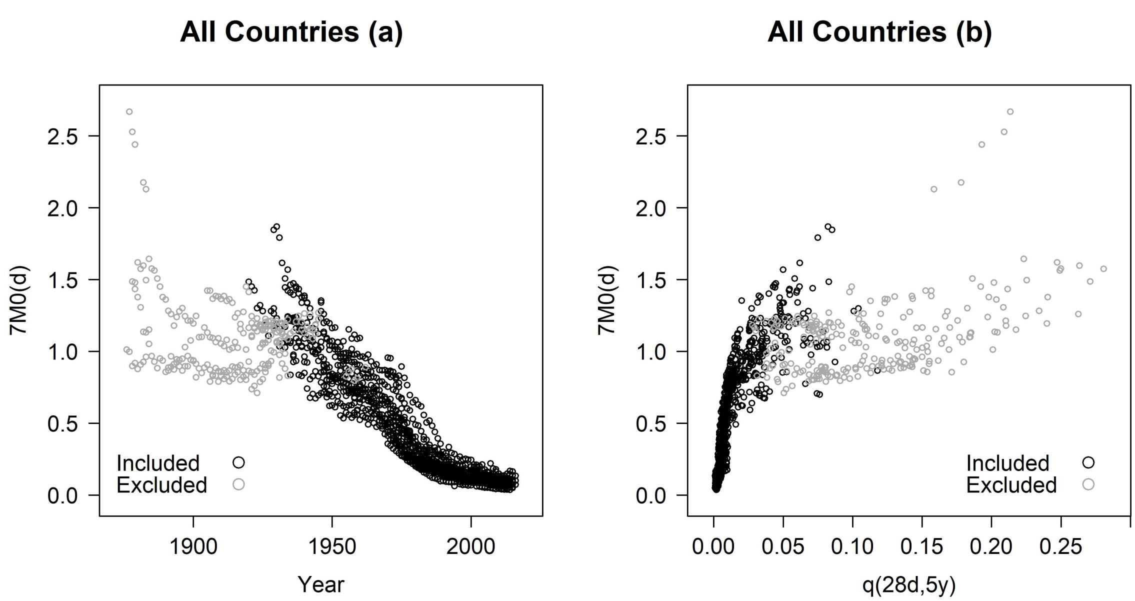
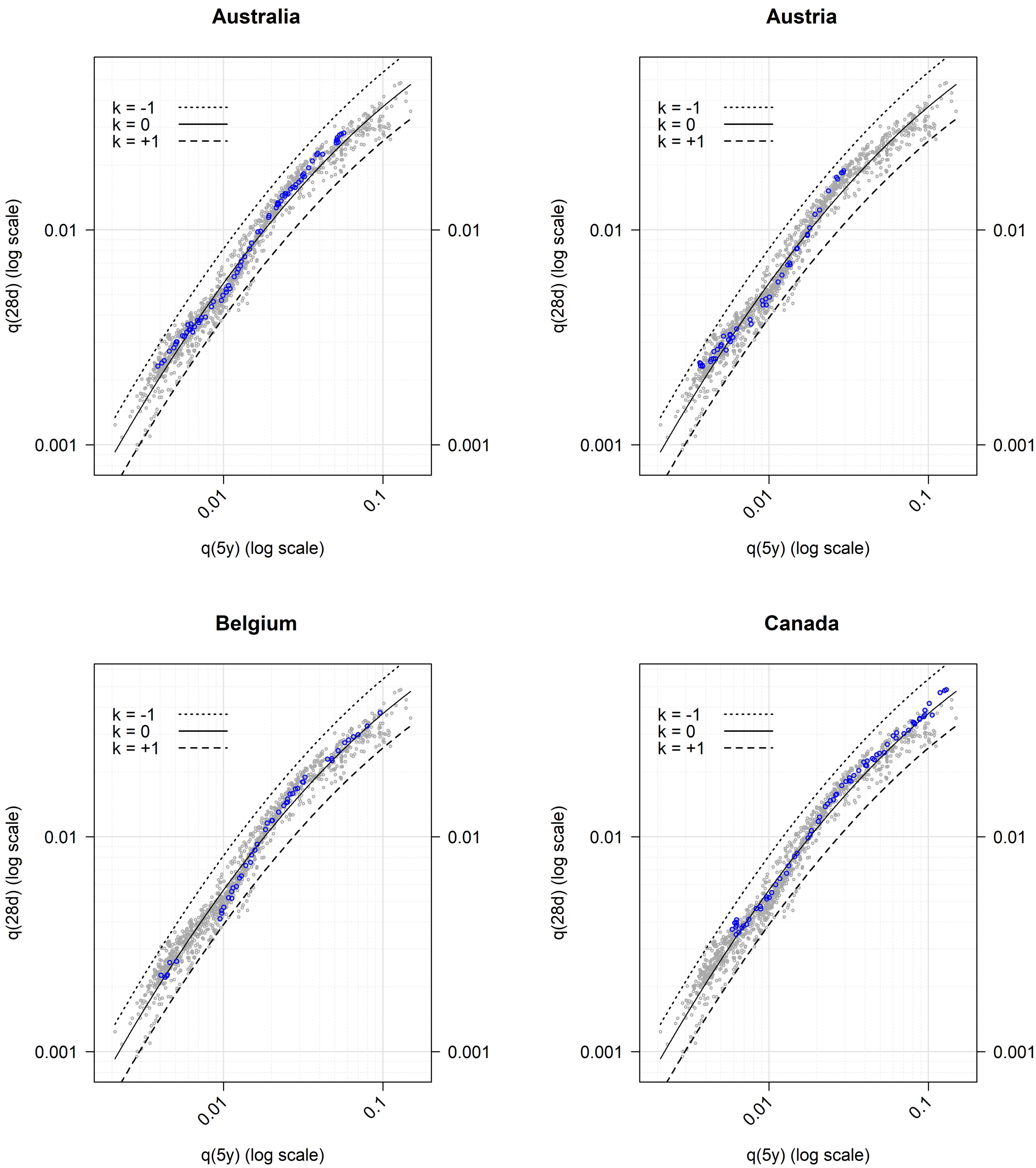
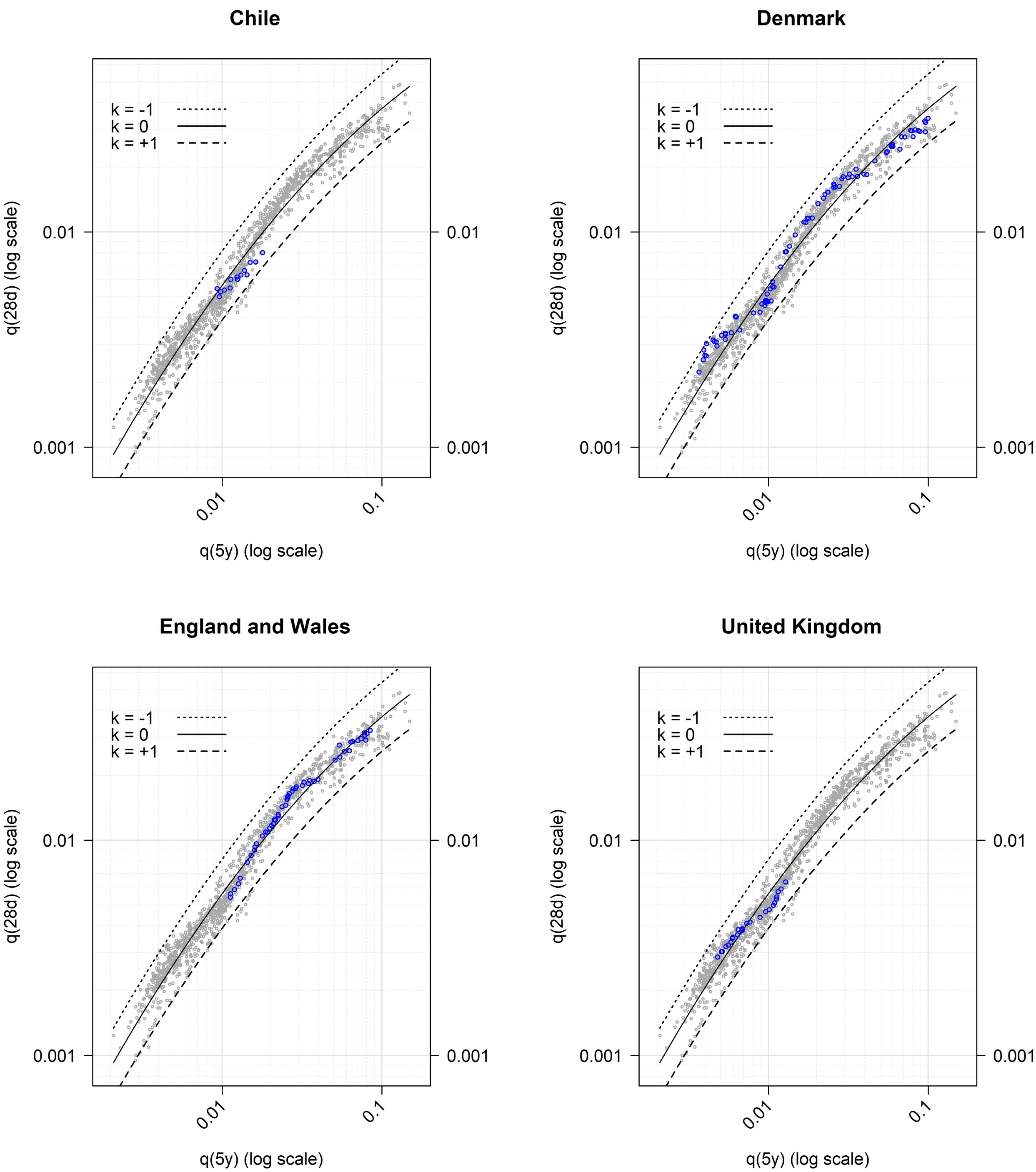
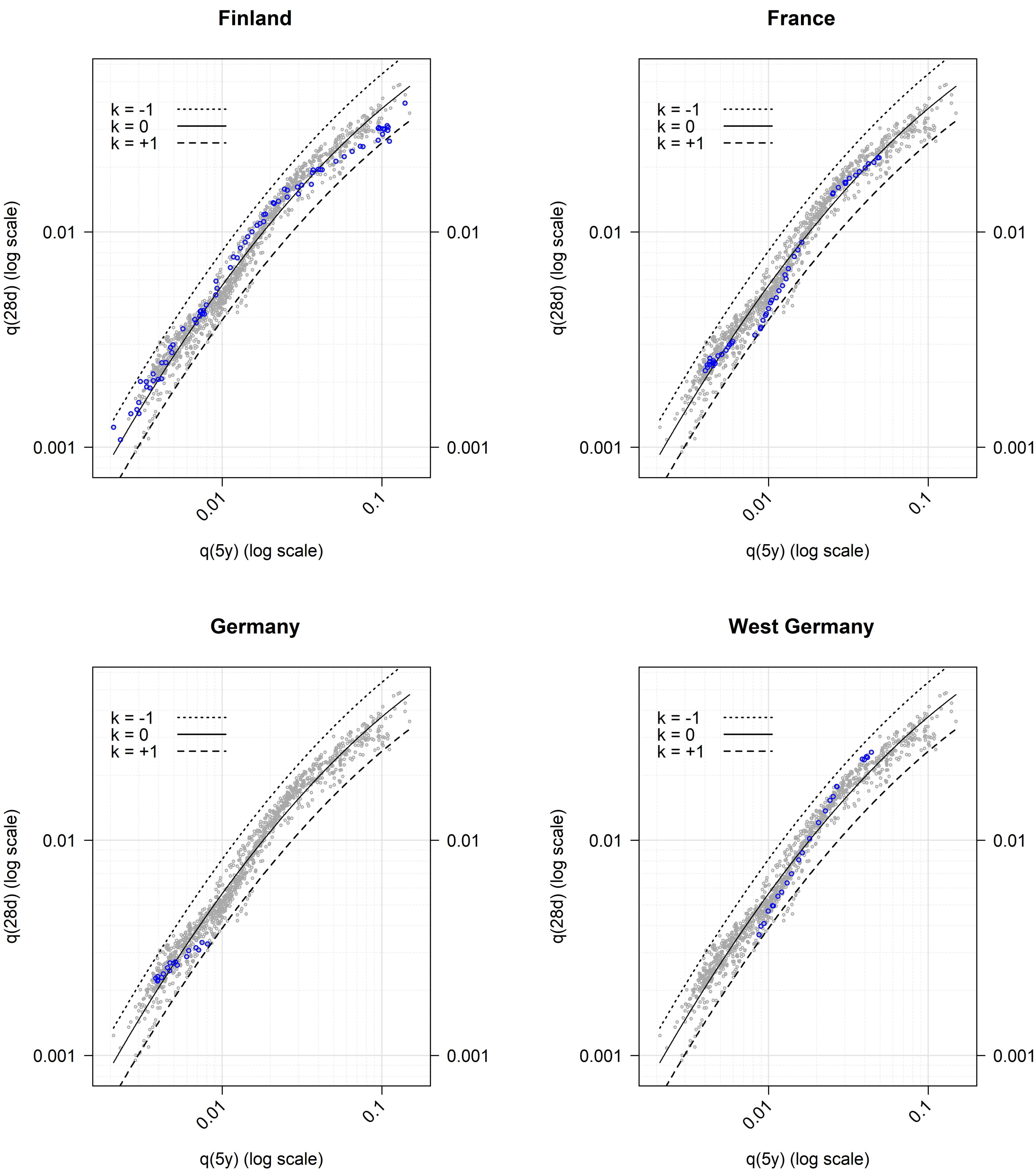
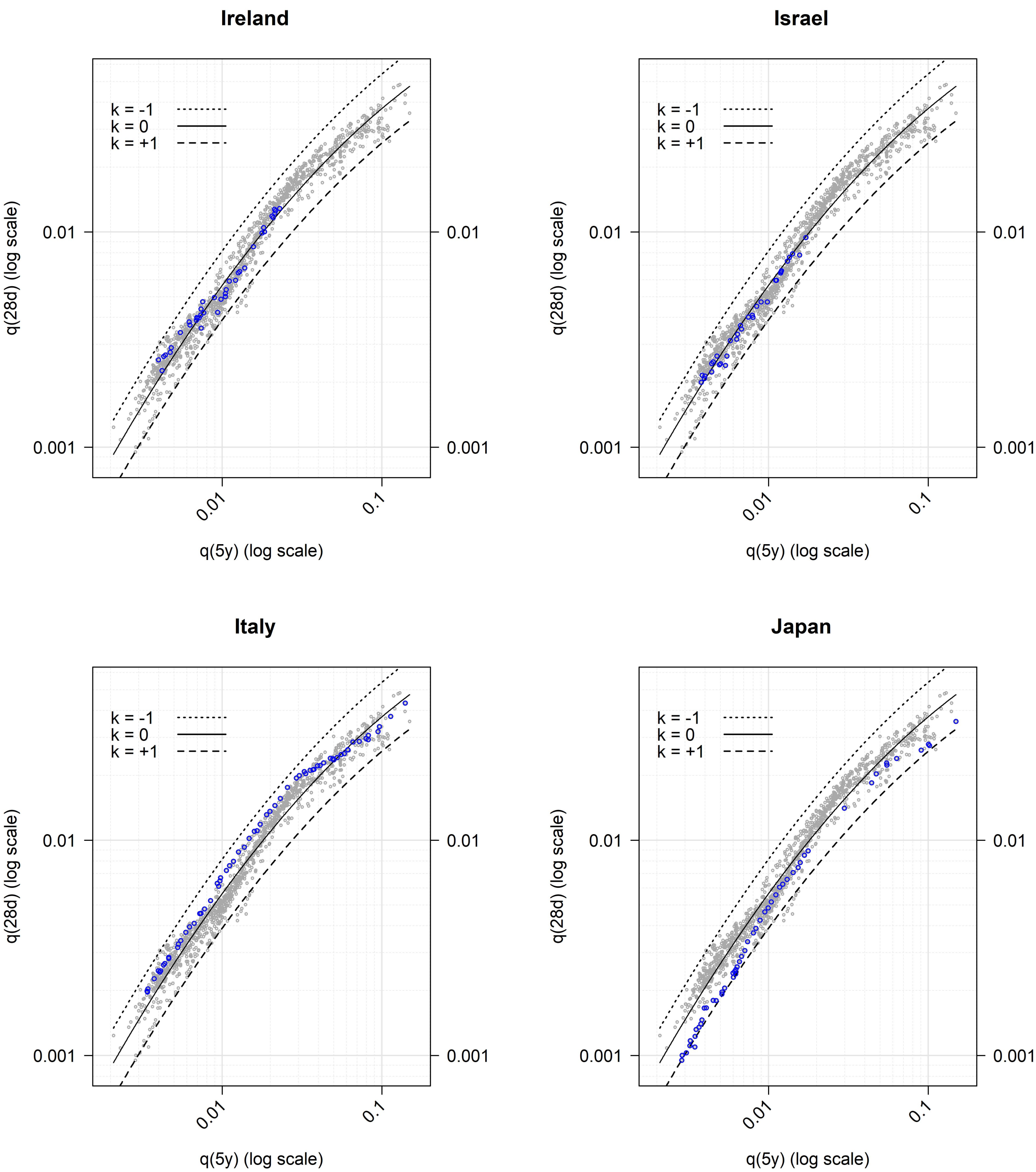


Figure SM2-3: Country-specific relationship between q(28d) and q(5y) with observed values in the final U5MD and values predicted using the log-quadratic model with k=0, ±1, both sexes combined

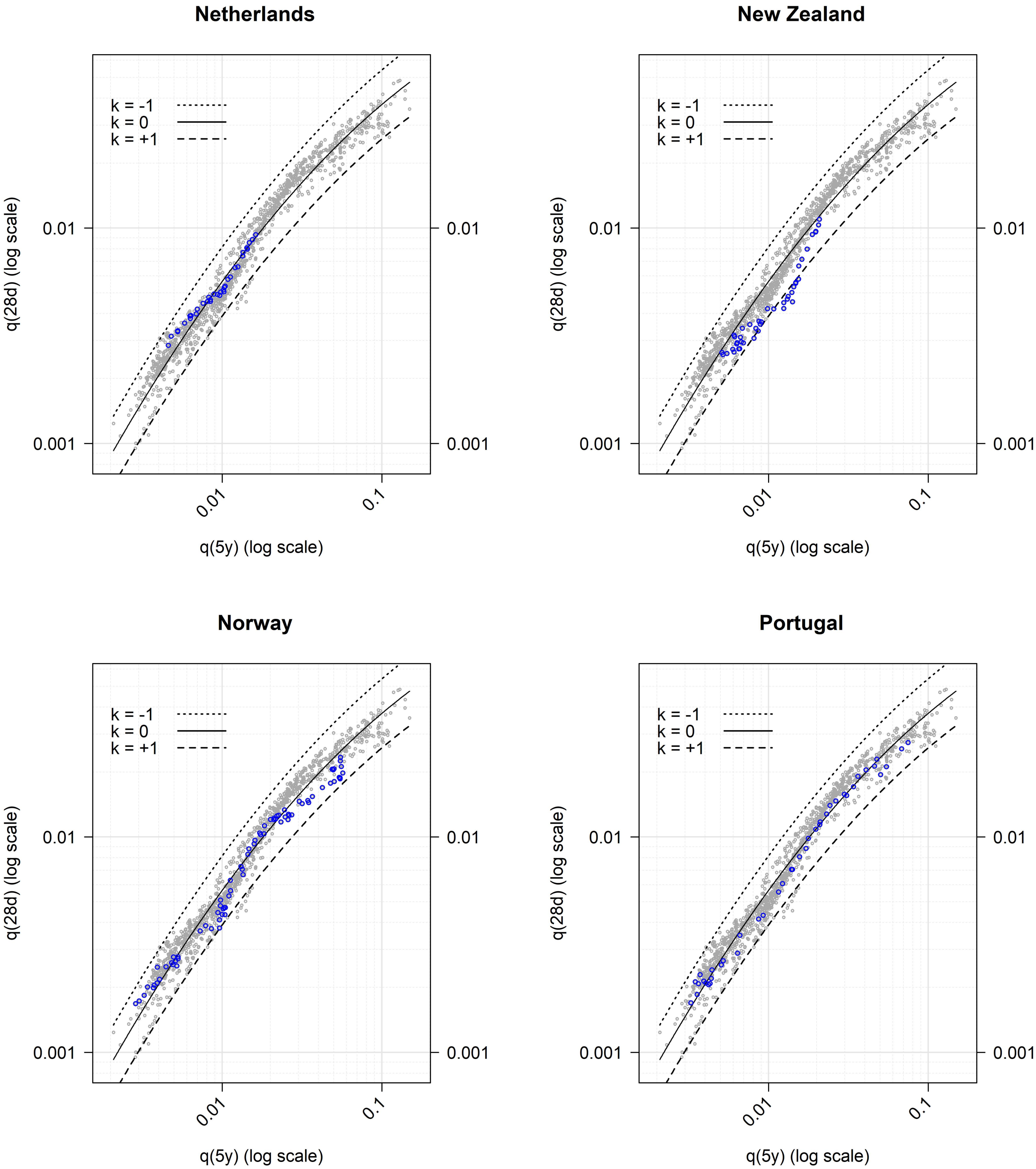


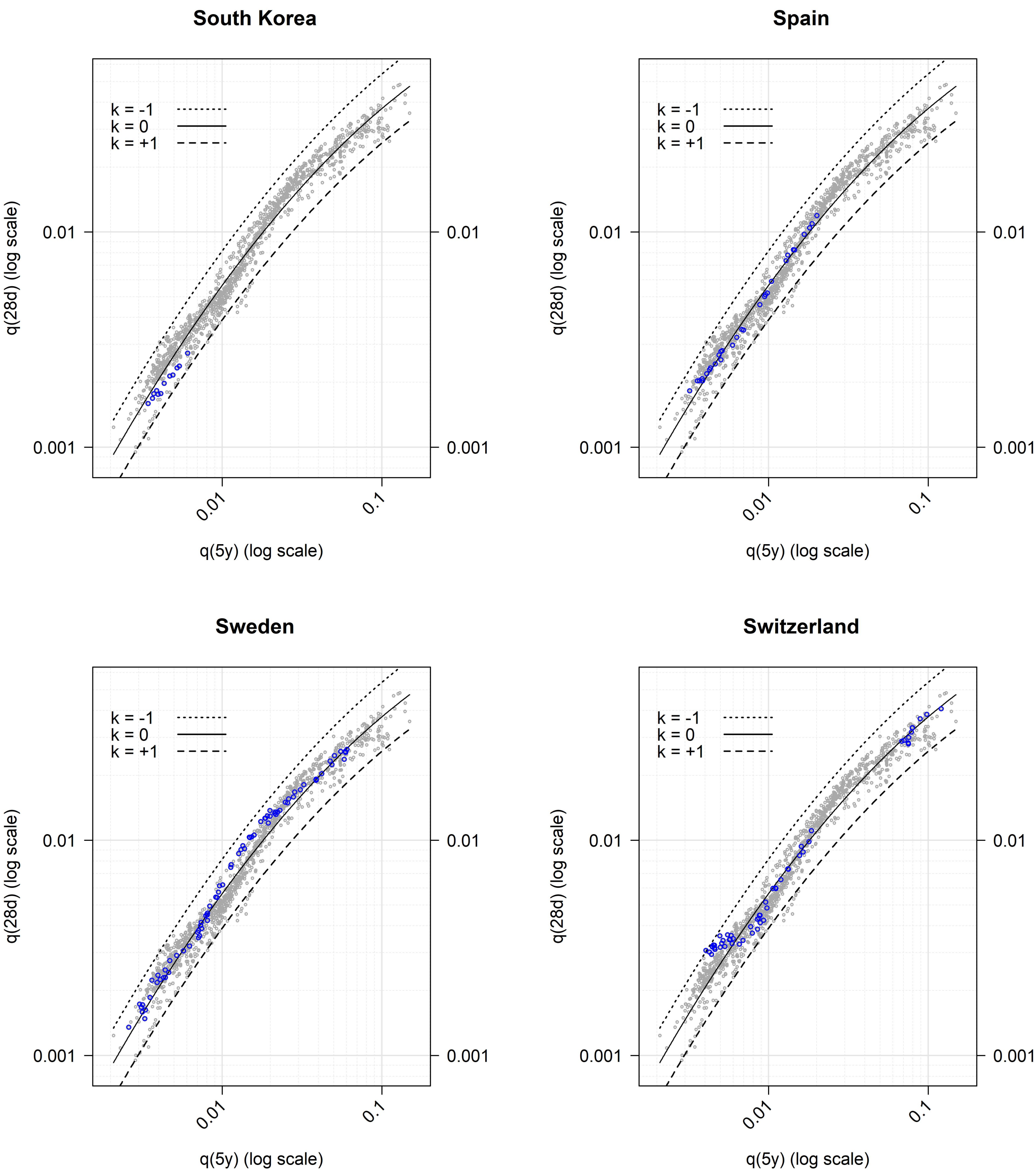

Chile
Denmark
England and Wales
United Kingdom
k = -1
k = 0
k = +1
q(28d) (log scale)
q(5y) (log scale)
0.01
0.001
0.1

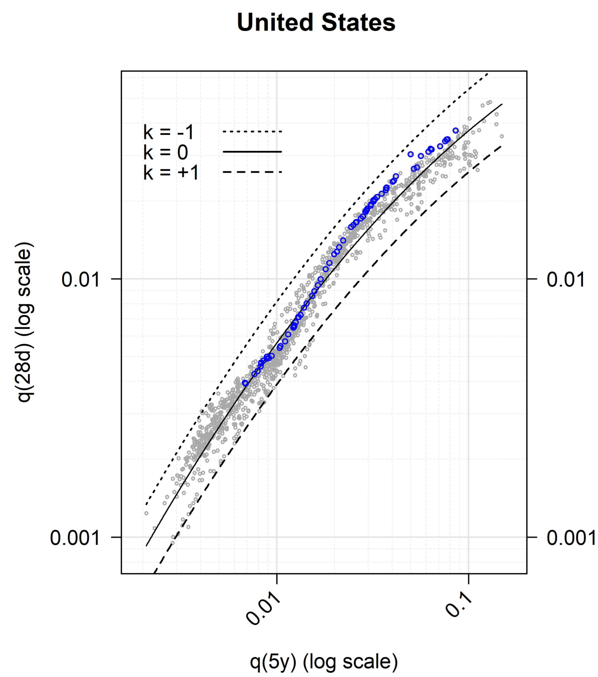

Finland
France
Germany
West Germany
k = -1
k = 0
k = +1
q(28d) (log scale)
q(5y) (log scale)
0.01
0.001
0.1



Ireland
Israel
Italy
Japan
k = -1
k = 0
k = +1
q(28d) (log scale)
q(5y) (log scale)
0.01
0.001
0.1


Netherlands
New Zealand
Norway
Portugal
k = -1
k = 0
k = +1
q(28d) (log scale)
q(5y) (log scale)
0.01
0.001
0.1


South Korea
Spain
Sweden
Switzerland
k = -1
k = 0
k = +1
q(28d) (log scale)
q(5y) (log scale)
0.01
0.001
0.1



United States
k = -1
k = 0
k = +1
0.01
0.001
0.01
0.001
0.01
0.1
q(28d) (log scale)
q(5y) (log scale)