Pension Systems / Demography & Mortality

Lecture notes: Demography – part II

University of Copenhagen, Autumn 2021

Snorre Jallbjørn & Søren F. Jarner

Recap – The model by Lee & Carter (1992)

- The death rates are modelled by a bilinear regression
  - $log\big(m(x,t)\big) = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t}$
    - α and β are age-dependent parameters, and κ is a time-varying “mortality index”.
    - $\varepsilon_{x,t}$ are assumed independent, mean-zero and with common variance
  - The model is over-parameterized: Given $\{\alpha_x, \beta_x, \kappa_t\}$ the triple $\left\{\alpha_x - \beta_x c,\ \frac{\beta_x}{d},\ d(\kappa_t + c)\right\}$ gives identical fit
    - Identification is typically ensured by the constraints $\sum_x \beta_x = 1$, and $\sum_t \kappa_t = 0$

- In the original paper, the model is fitted by OLS, i.e. minimization of $\sum_{t,x}(\log m(x,t) - \alpha_x - \beta_x \kappa_t)^2$
  - $\hat{\alpha}_x =$ time average of $\log m(x,t)$; $\hat{\beta}_x$ and $\hat{\kappa}_t$ are the first components of a SVD of the matrix $\{\log m(x,t) - \hat{\alpha}_x\}_{x,t}$

- $\{\kappa_t\}$ is typically modelled as a random walk with drift: $\kappa_t = \kappa_{t-1} + \theta + \varepsilon_t$, with iid $\varepsilon_t \sim N(0, \sigma^2)$
  - For $h \geq 1$ we have $\kappa_{T+h} | \kappa_T \sim N\big(\kappa_T + h\hat{\theta}, h\hat{\sigma}^2\big)$
  - Median forecast is given by: $\bar{\mu}(x, T+h) = \exp\big(\hat{\alpha}_x + \hat{\beta}_x \kappa_{T+h}\big) = \exp\big(\hat{\alpha}_x + \hat{\beta}_x \kappa_T + \hat{\beta}_x h\hat{\theta}\big) = \hat{\mu}(x,T)\exp\big(h\hat{\beta}_x\hat{\theta}\big)$

Recap – The balancing equation

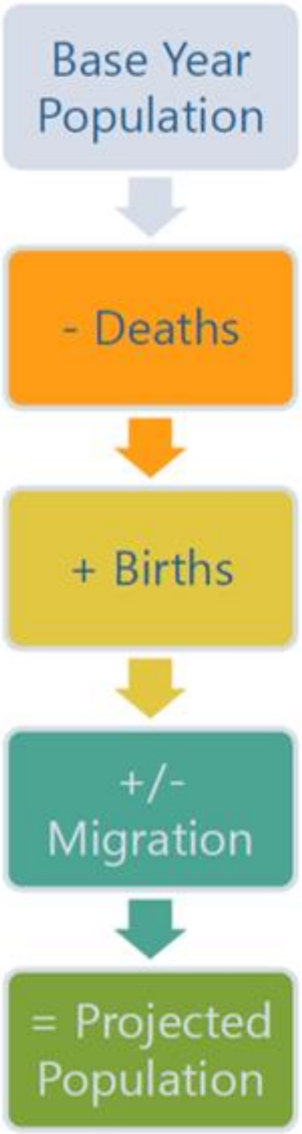
- The "balancing equation" of population change: $N_1 = N_0 + (B - D) + (I - E)$
  - N_0 and N_1 are the population sizes at the beginning and end of a period of interest, e.g. a year
  - B is the number of births in the period
  - D is the number of deaths in the period

  $B - D$ is the *natural increase*

  - I is the number of immigrations over the period
  - E is the number of emigrations over the period

  $I - E$ is the net migration

- Rearranging the balancing equation and divide by person-years lived in the period (PY)
  - $\frac{N_1 - N_0}{PY} = \frac{B}{PY} - \frac{D}{PY} + \frac{I}{PY} - \frac{E}{PY}$
  - Crude growth rate = crude birth rate – crude death rate + crude rate of net migration
    - 'Crude' refers to the fact that the rates are relative to the entire population
    - In contrast, an age-specific rate, e.g. of mortality, is a rate relative to a specific age-group



#1

Fertility

Fertility

- Definition: ”*The increment process by which members of a population produce live births*” - Preston et al. (2001) p. 92.

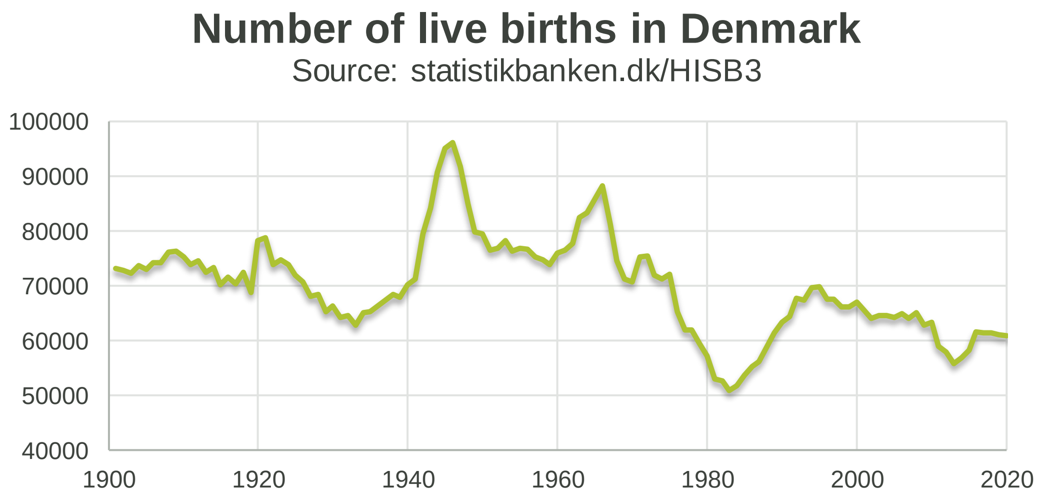
- Modelling and projecting fertility (rates) is a bit different from mortality
  - Fertility (usually) involves two individuals of opposite sexes
  - ”Risk” of producing birth is not necessarily strictly positive at all ages
    - The fertile age range (reproductive span) is $[\alpha, \beta]$, where α and β denote minimum and maximum ages of childbearing, respectively.
  - Fertility is highly correlated with e.g. marriage and previous number of births (parity)
  - Conceptual difference: Giving birth is a repeatable event

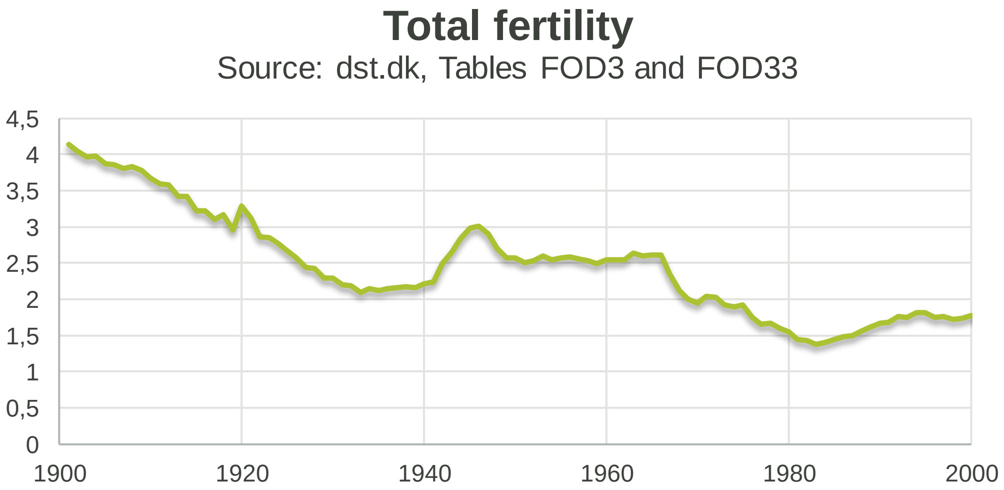
1. Danish fertility – Historic evolution

- **The size of birth cohorts has varied considerably through the 20th century**
  - The largest cohort (1946) has more than 96.000 births …
  - … and is almost twice as large as the smallest cohort (1983) with fewer than 51.000 births

- **Variation is caused by**

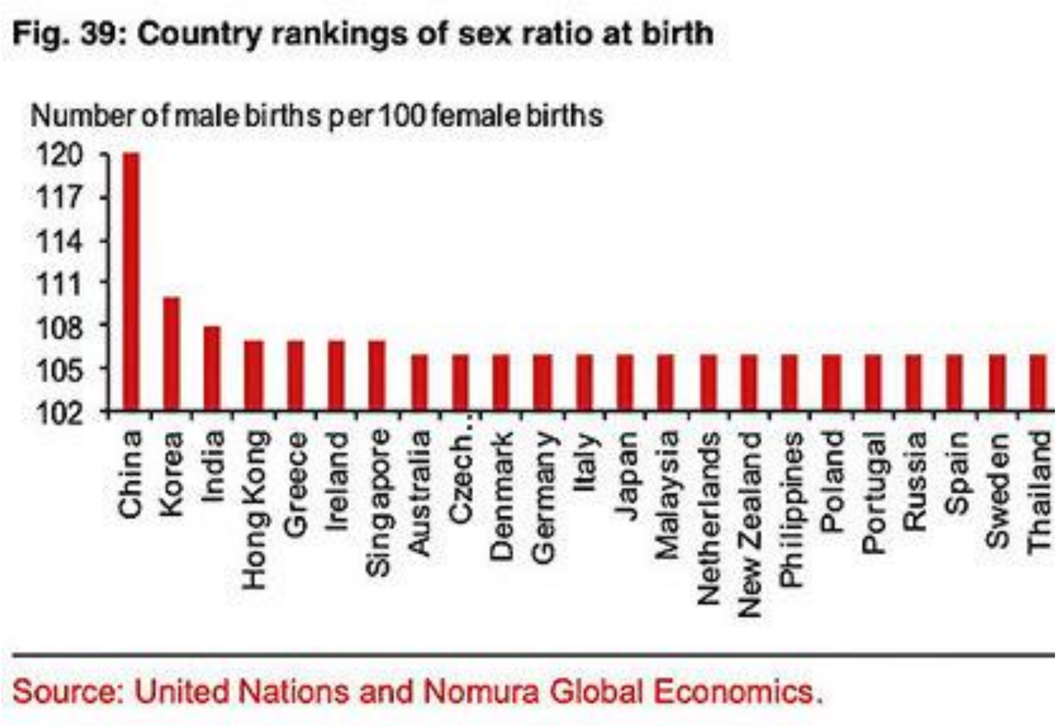
1. Variation in the number of fertile women
2. Variation in total fertility
   - Total fertility has decreased from **4.14** children per woman in 1901 to **1.67** per woman in 2020

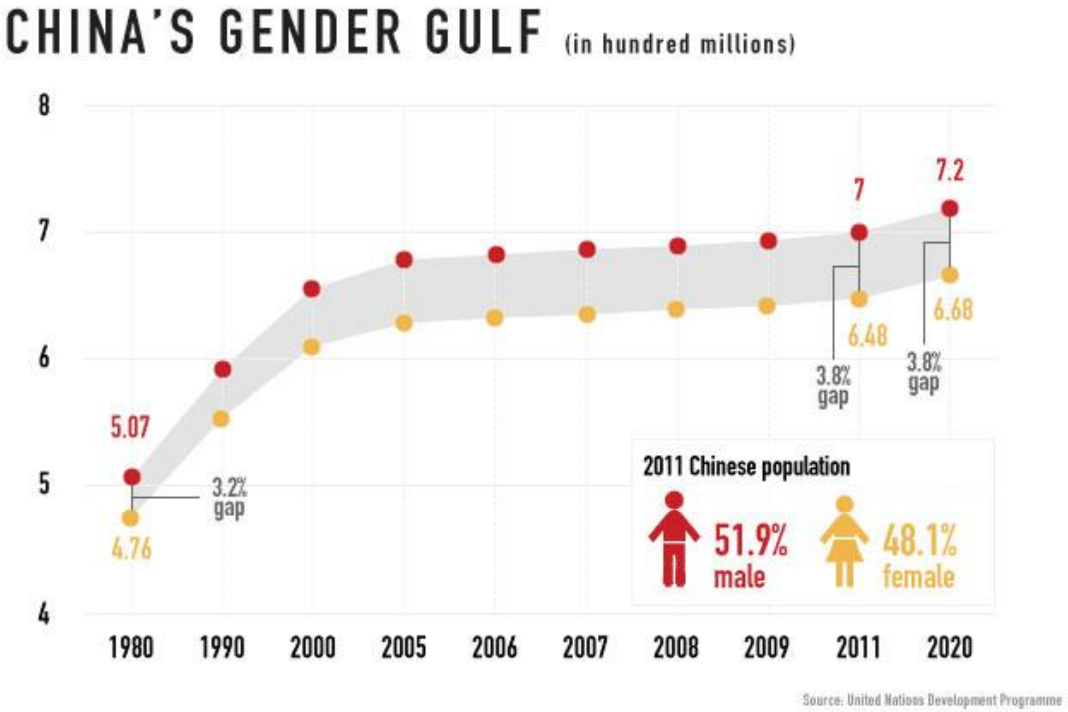




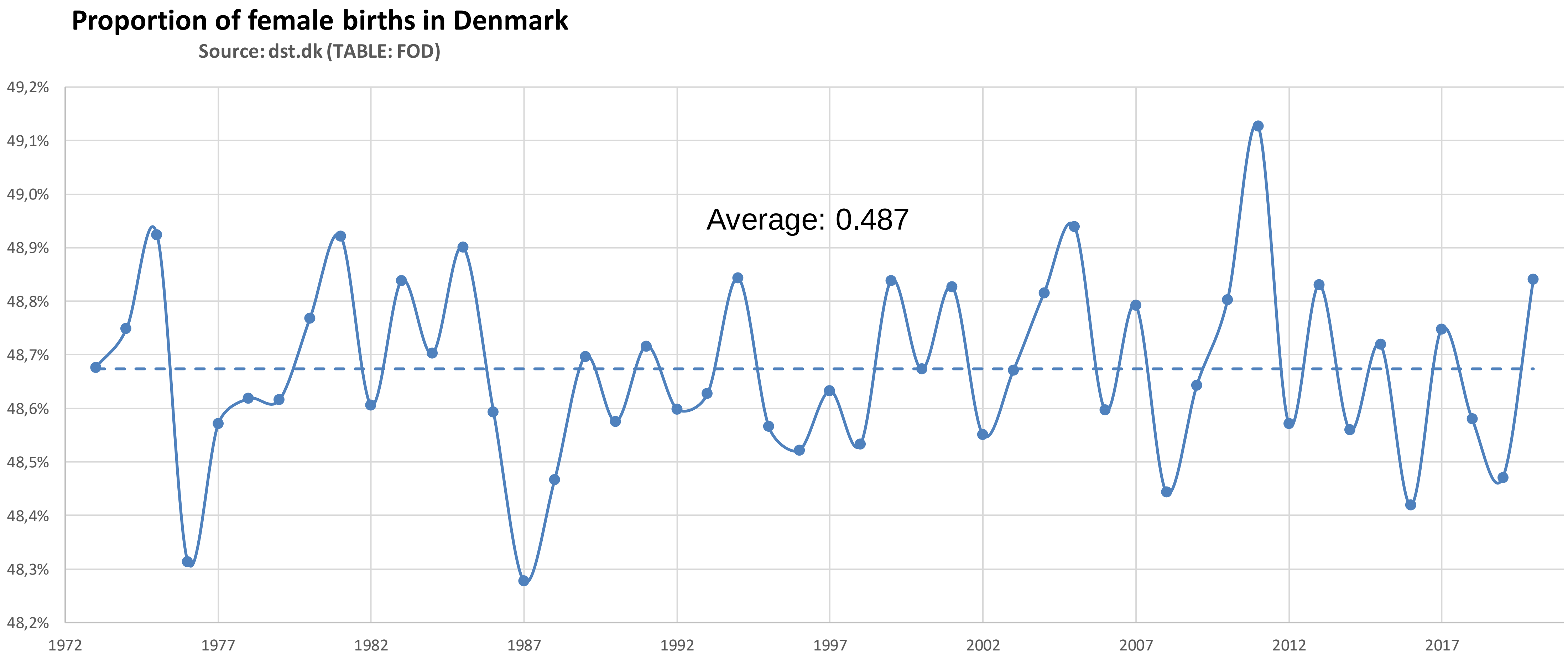
1. A short digression on two-sex modelling

- Our projection methodology is a so-called “female dominant” model
  - Births are independent of the male population, which works well as long as the sex ratio is stable
  - There exist two-sex models for fertility explicitly modelling the joint dependence on both sexes

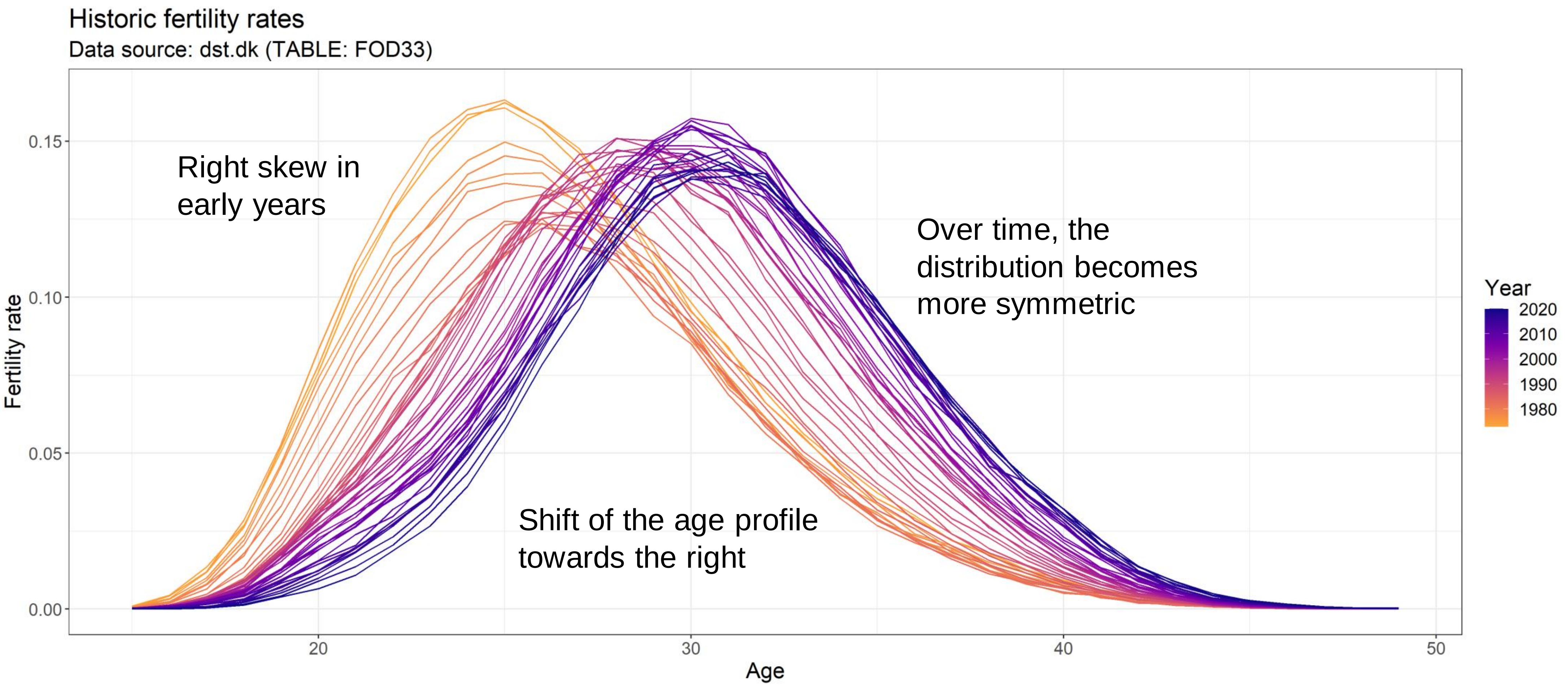




1. Danish fertility – Proportion of female births



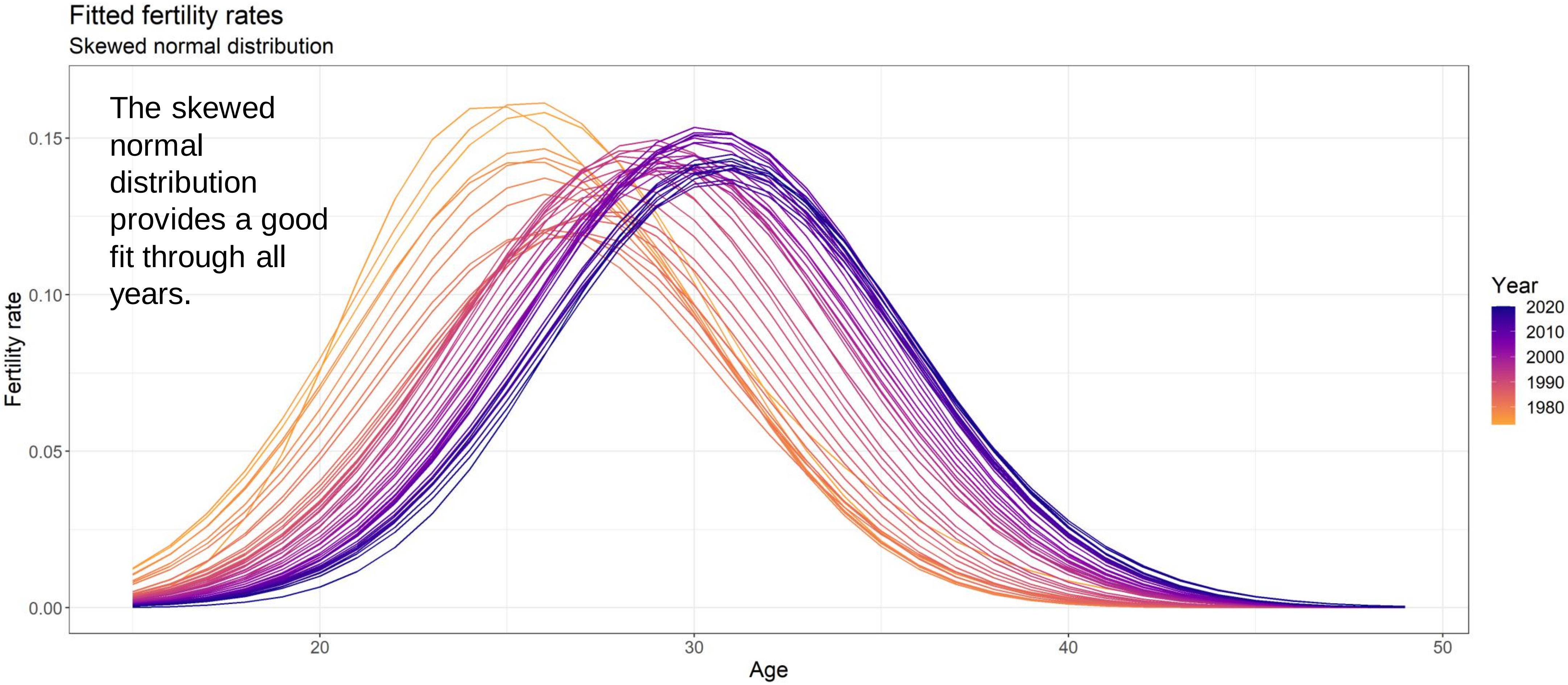
1. Danish fertility – Historic fertility rates



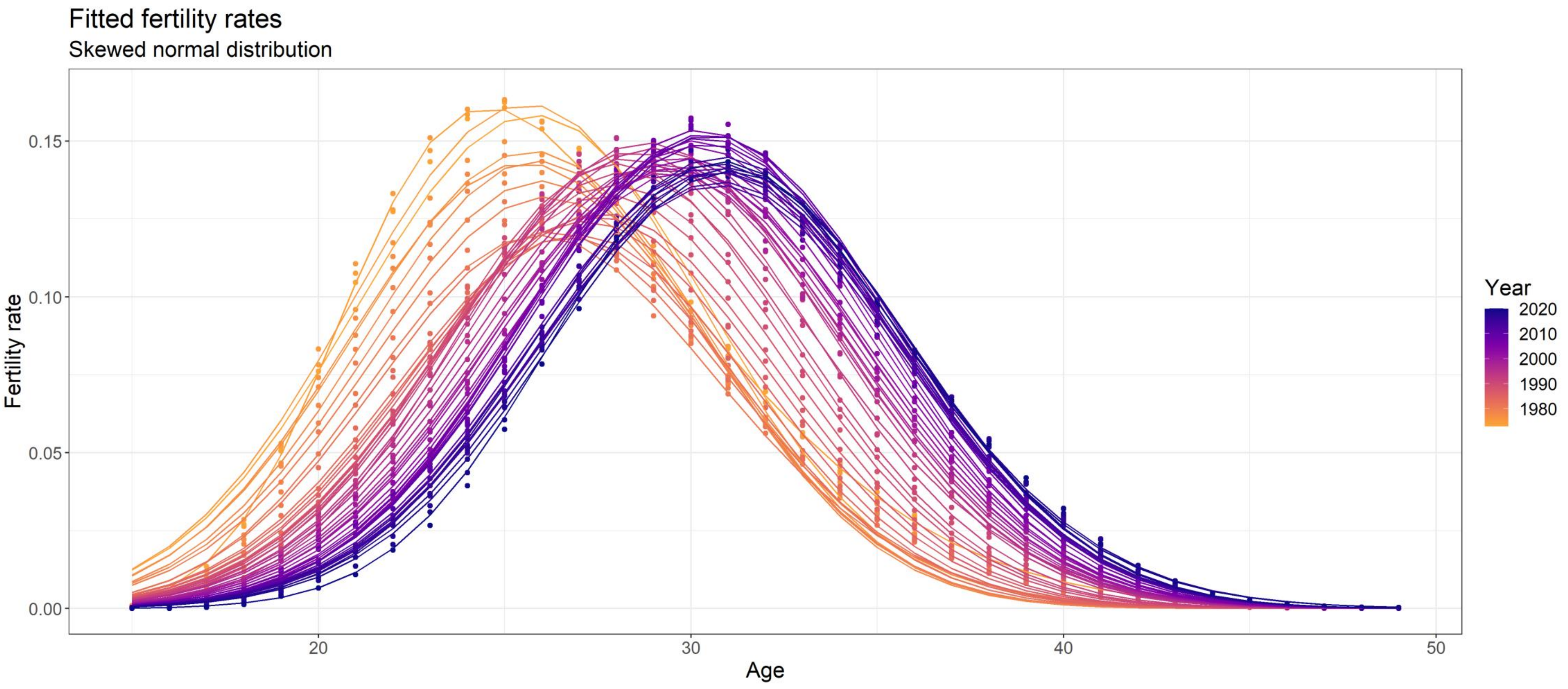
Tip: You can use the $plotdatFer-method (with a fertility-object) to create this plot.

1. Danish fertility – Fit of fertility rates (Skewed normal)

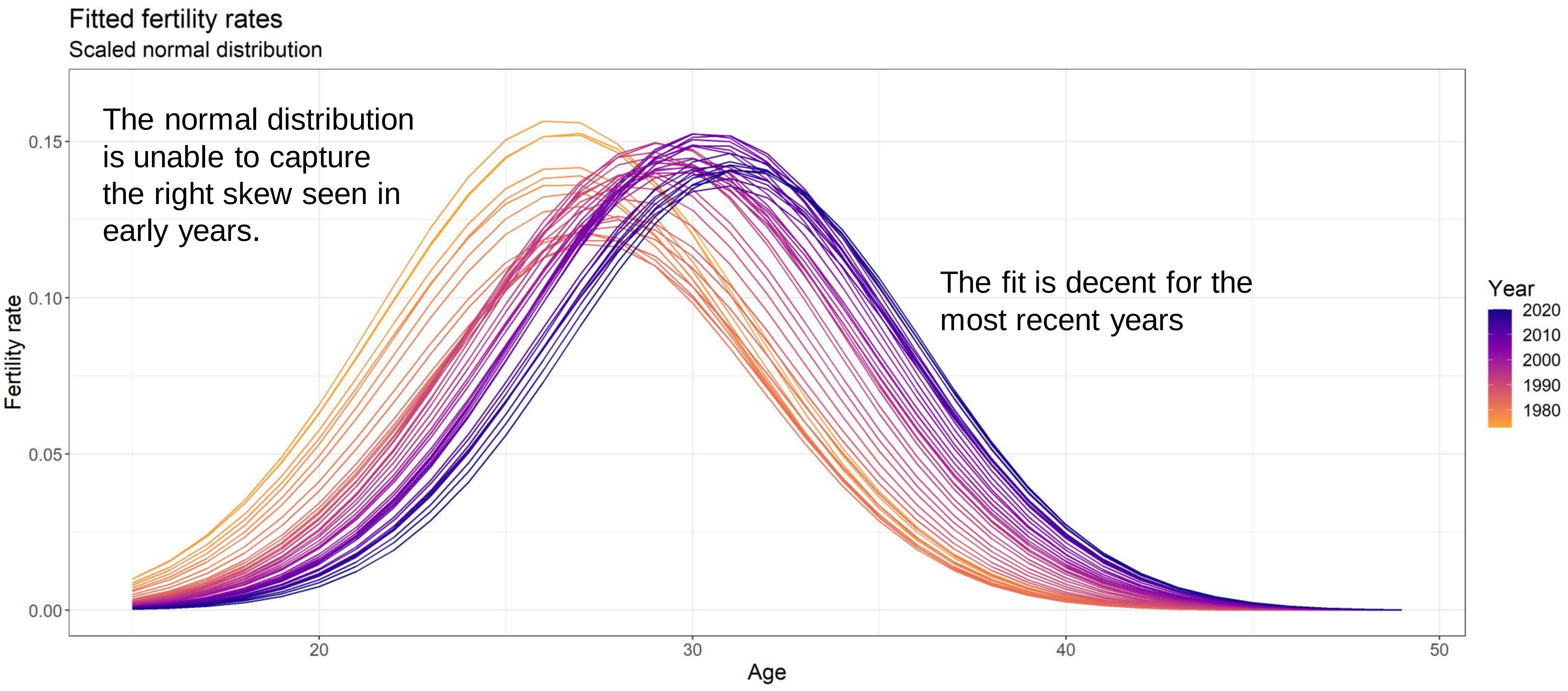
The skewed normal distribution provides a good fit through all years.



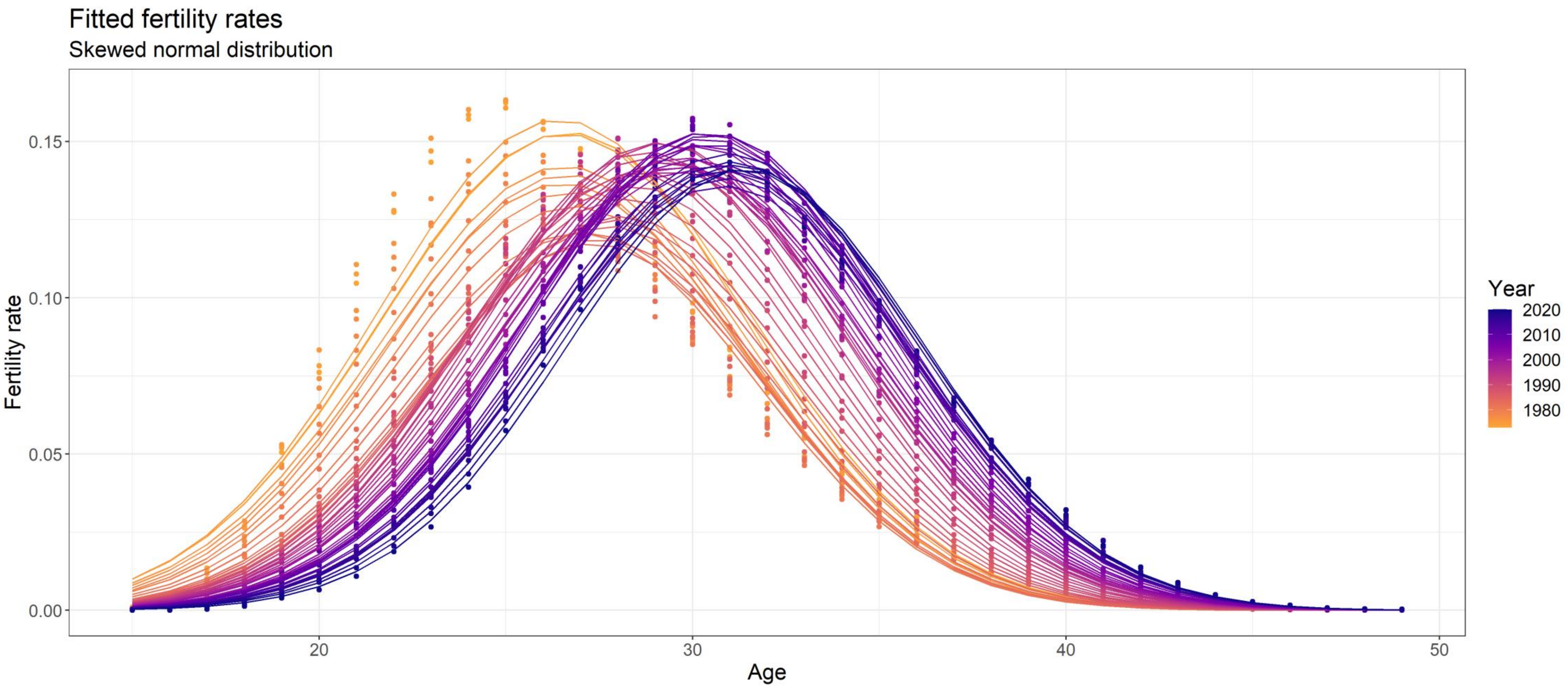
1. Danish fertility – Fit of fertility rates (Skewed normal)



1. Danish fertility – Fit of fertility rates (Scaled normal)

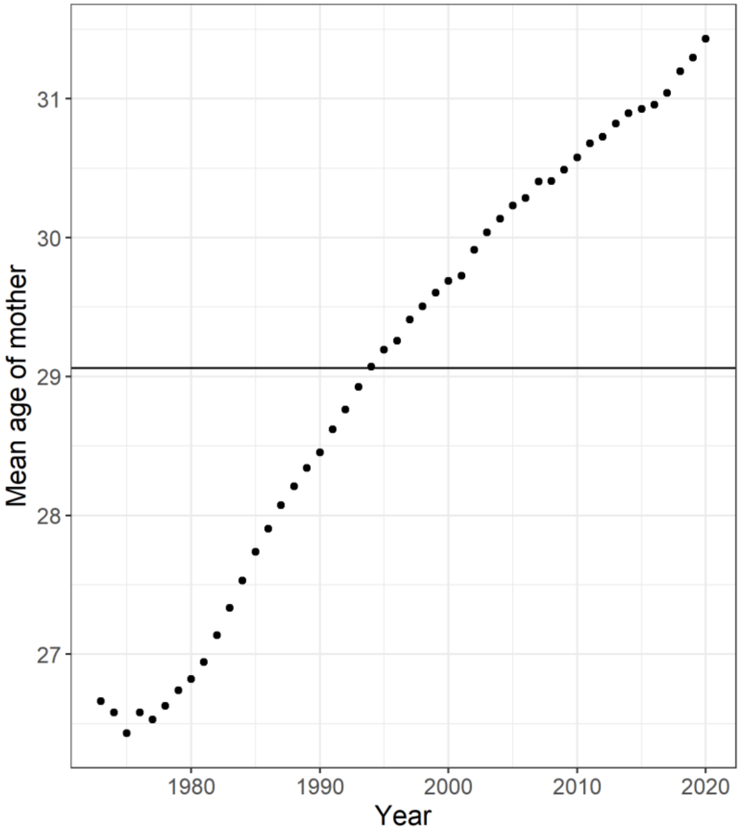


1. Danish fertility – Fit of fertility rates (Scaled normal)

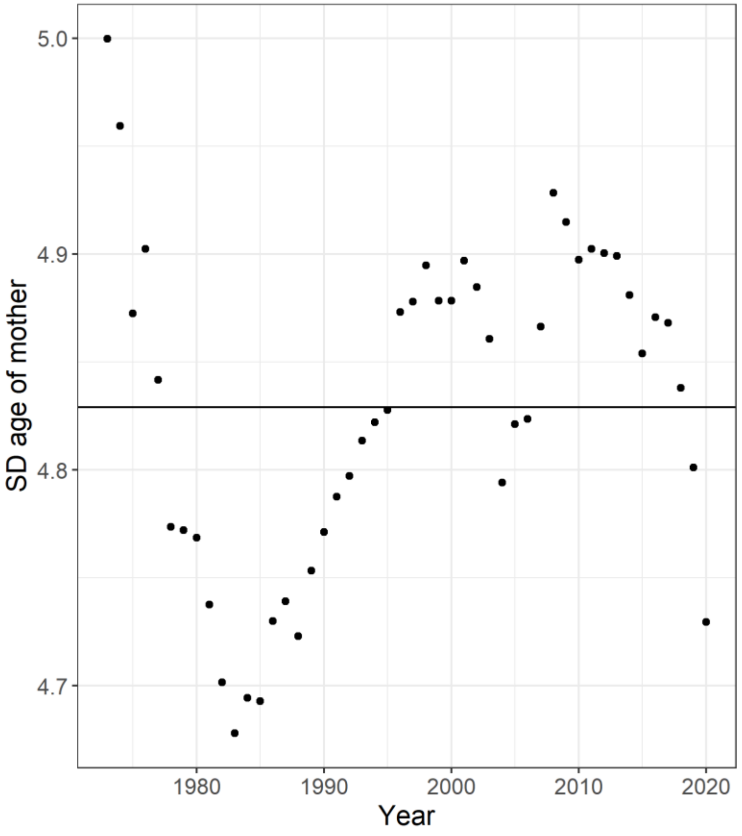


1. Danish fertility – historic summary statistics

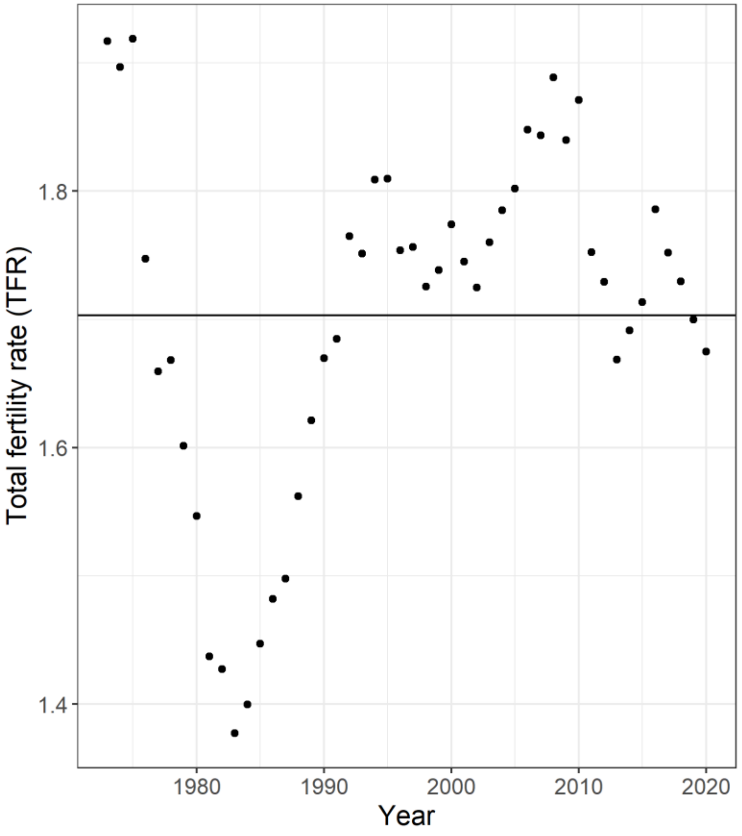
Mean age of mother slightly above 29 years, but increasing over time – how to project this?



Standard deviation of age of mother has been in a fairly narrow band around 4.85 years

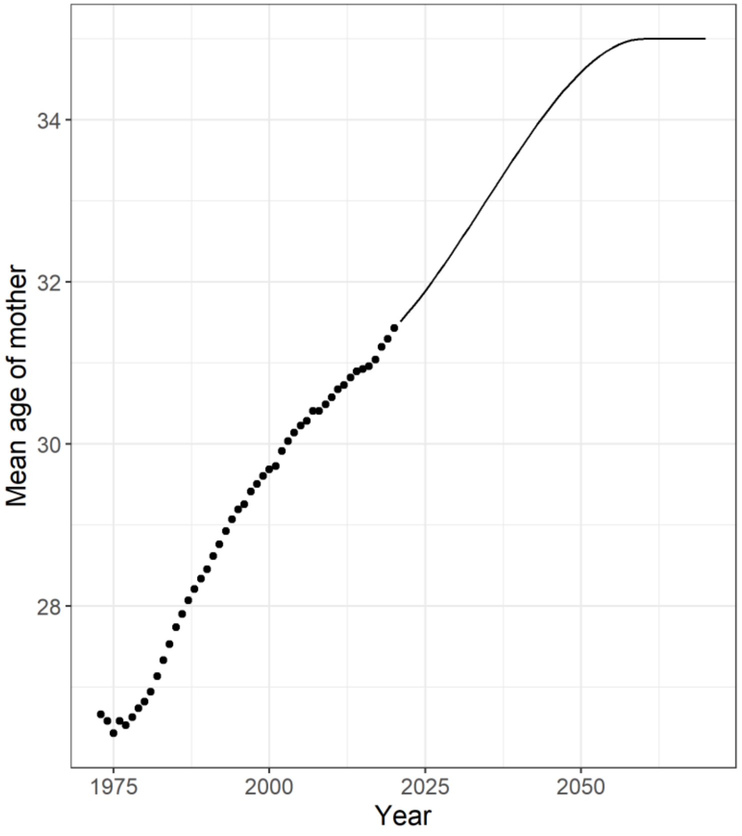


Total fertility fluctuating, but well below replacement level of 2.1 throughout the period

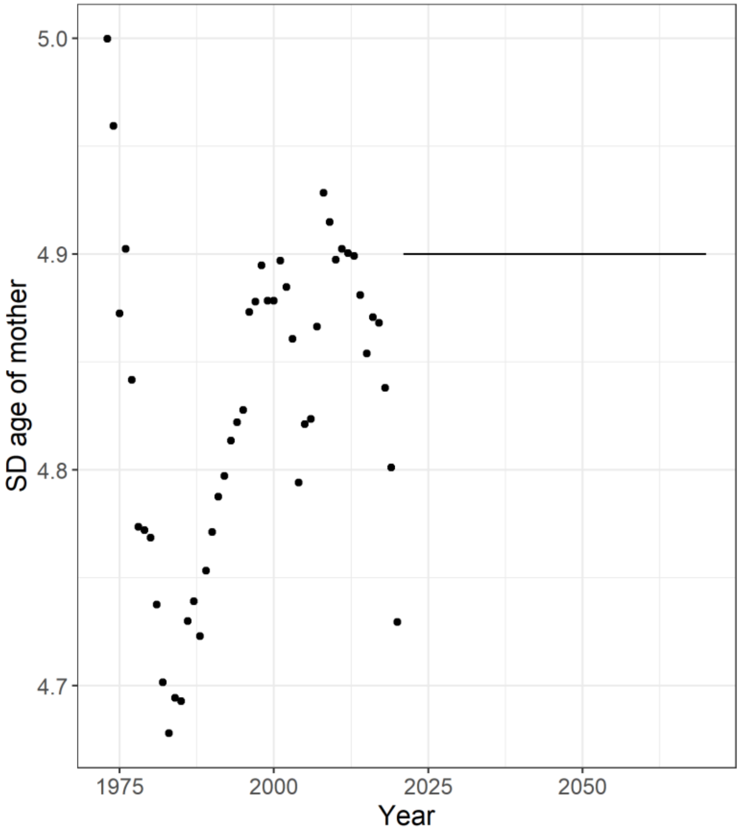


1. Danish fertility – historic summary statistics

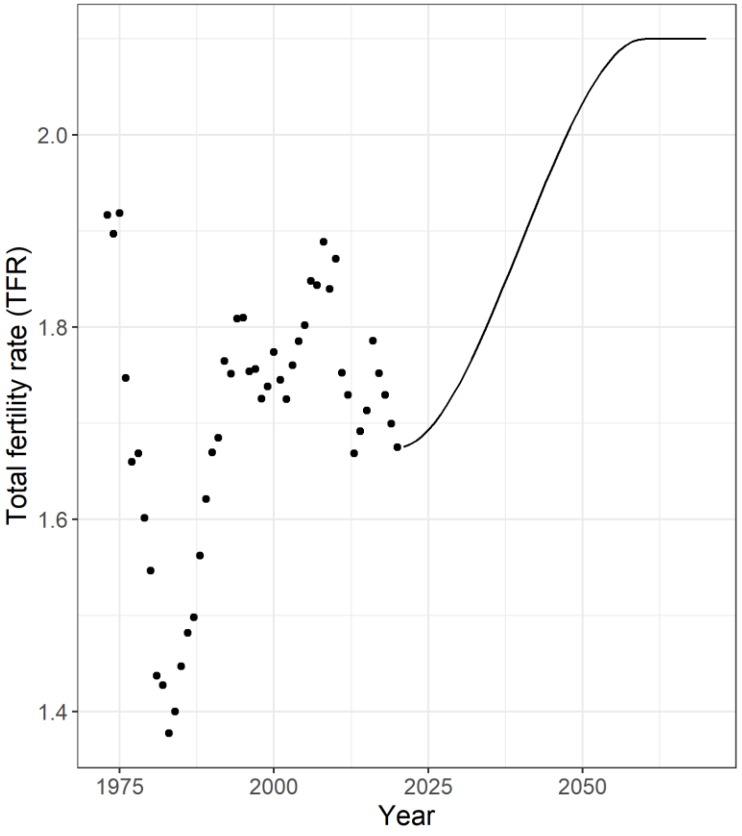
Mean age of mother increasing from current age to age 35 in 2060; Initial slope of 2 years pr. 25 years



Standard deviation of age of mother is kept constant at 4.9 years



Total fertility is increasing to replacement level of 2.1 in 2060; Initial slope of 0



#2

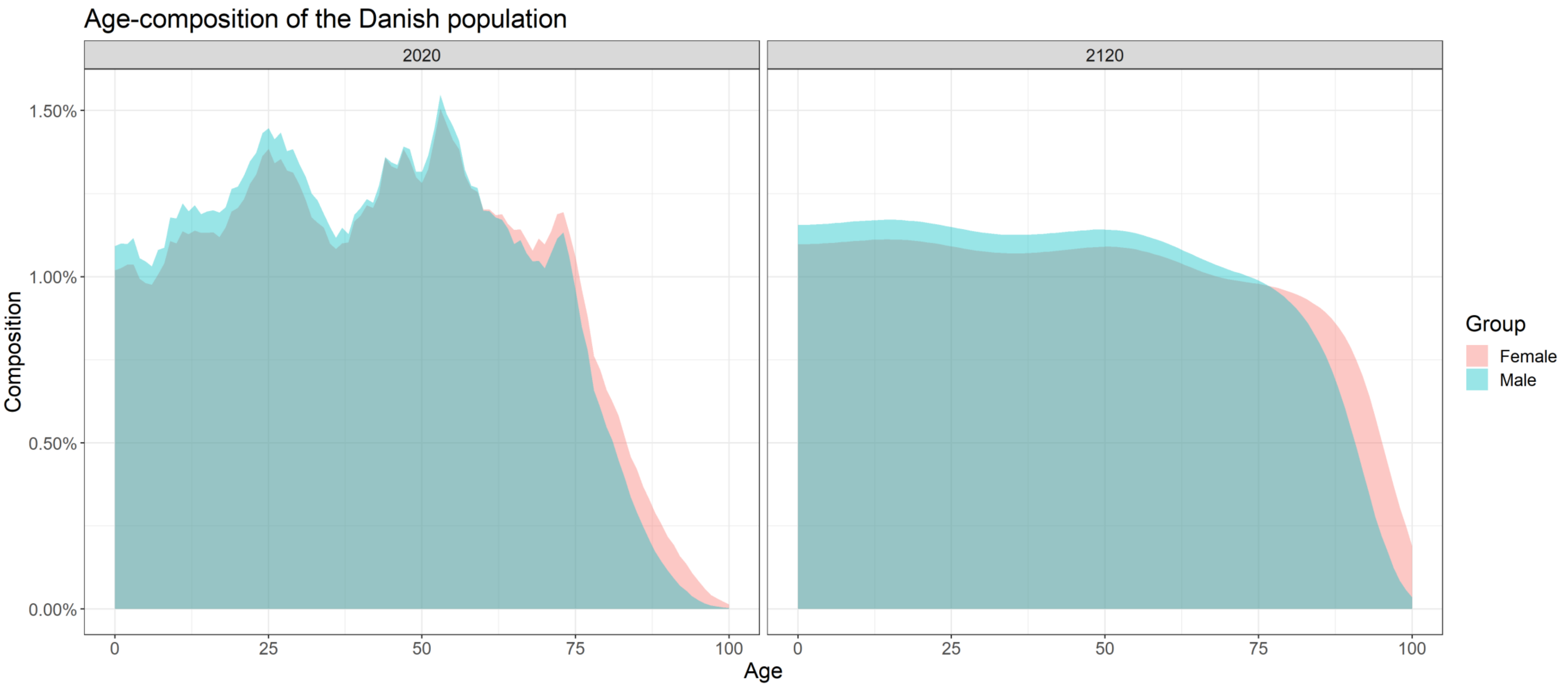
Population projection

2. The Cohort Component Method

- The method usually used for population projections is called the “cohort component” method
  - It consists of segmenting the population into subgroups differentially exposed to the “risks” of fertility, mortality, and migration – and computing the changes over time for each group
  - Virtually all demographic projections segregate by sex and age
  - Depending on the application, we might also differentiate between location, education, income …

- For each time step in the projection, the method consists of three steps:
  1. For each subgroup, compute the number still alive at the beginning of the next interval
  2. Compute the number of births for each subgroup over the interval, add them across groups, and compute the number of those births that survive to the beginning of the next interval
  3. For each subgroup, add immigrants (and surviving births) and subtract emigrants over the interval
  - *If the projection has more characteristics than sex and age, a simple approach often used is to assume that newborns belong to the same segment as the mother*

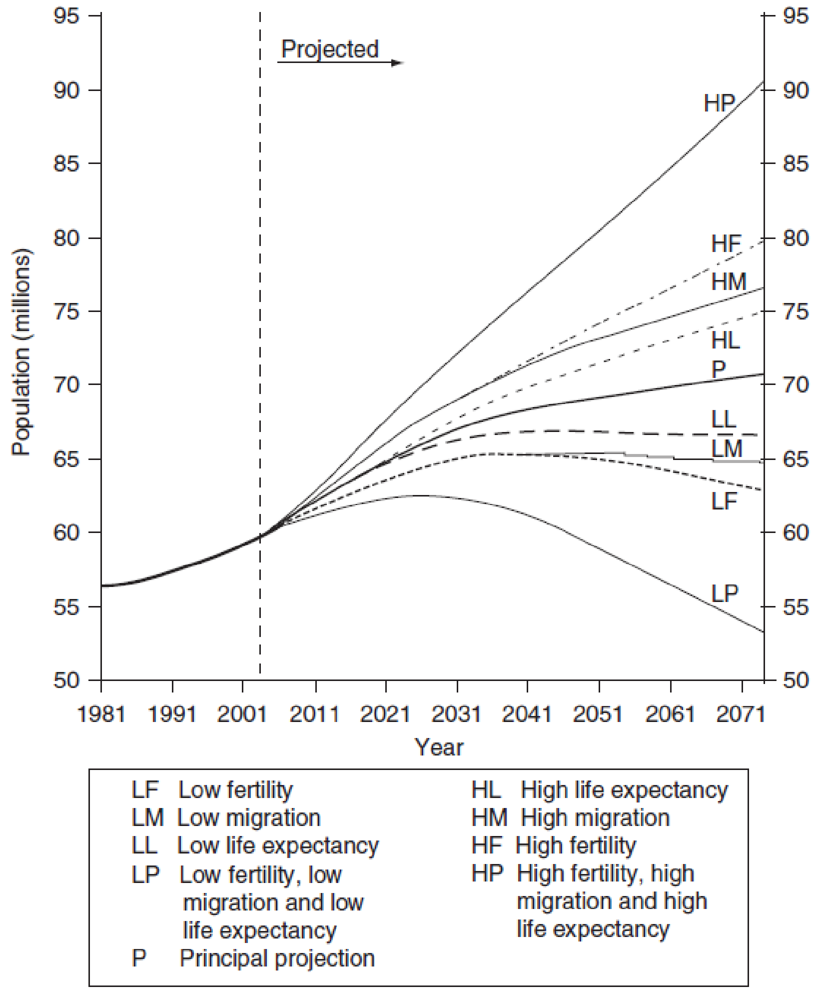
The Danish population



Tip: You can use the $plotdat-method with a pop-object to create this plot.
The geom used is geom_area(position=identity, alpha=0.4).

2. Deterministic vs. probabilistic projections

Main scenario (‘principal’) with stressed variants



Median scenario with probabilistic uncertainty

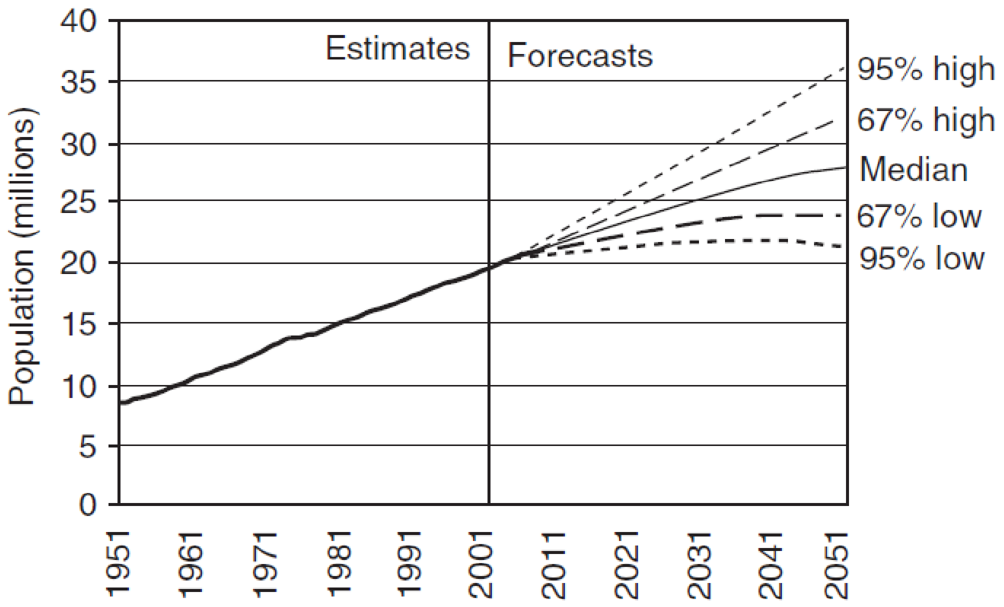


Figure 3 The probability distribution of Australia’s future population. Used with kind permission of Wilson, T. and Bell, M. (2004). Australia’s uncertain demographic future. *Demographic Research* 11(8), 195–234. Online at: http://www.demographic-research.org/Volumes/Vol11/8/, Figure 12.

3. The stable equivalent population

- The stable equivalent population of a given population is the population that would eventually arise, if current (or otherwise specified) mortality and fertility rates were to continue indefinitely

The Danish population 2020 and its stable equivalent ($r \approx -0.65\%$)

Projection to 2120 using 2020 rates