UNIVERSITÀ DEGLI STUDI DI PADOVA

Harnessing demographic skewness: insights and applications

Stefano Mazzuco
drawing from joint works with E. Aliverti, V. Canudas-Romo, D. Durante, B. Scarpa, M. Suhrcke, L. Zanotto

Skew 2026, Padova, January 8, 2026

Skewness in Demography?

- Do we need ~~Skew-symmetric~~ symmetry-modulated distributions in demography?
- Most demographic phenomena (fertility, mortality) are age-structured and often asymmetric.
- However skewness of fertility and mortality schedules is rarely considered
- Focus is mainly on:
  - Timing, with different indices: Mean age at childbearing, Modal age at death
  - Inequality, using measures of dispersion: standard errors, variance
- Adding a skewness parameter to models fitting fertility and mortality schedules might be useful
- ...but is there a demographic meaning of skewness?

Symmetrization of fertility?

Modelling fertility schedules

We start from fertility model representation proposed by Hoem (1981), where Age–Specific Fertility Rates may be written as

$$g(x; R, \theta_2, \ldots, \theta_r) = R \cdot h(x; \theta_2, \ldots, \theta_r) \tag{1}$$

- Note that model separates the “quantum” (given by parameter R) and “tempo” components – the pdf $h(\cdot)$.
- Therefore it can be profitably used to determine the effect of changing the timing of fertility schedules on the quantum parameter

Modelling fertility schedules

- Often used skewed distribution for $h(\cdot)$: Gamma, Hadwiger
- But in most recent years, many low fertility countries have witnessed an increasingly symmetric distribution of births
- Peristera & Kostaki propose a normal distribution with varying variance (before and after the mode)
- Mazzuco & Scarpa (2015) suggested a skew normal distribution

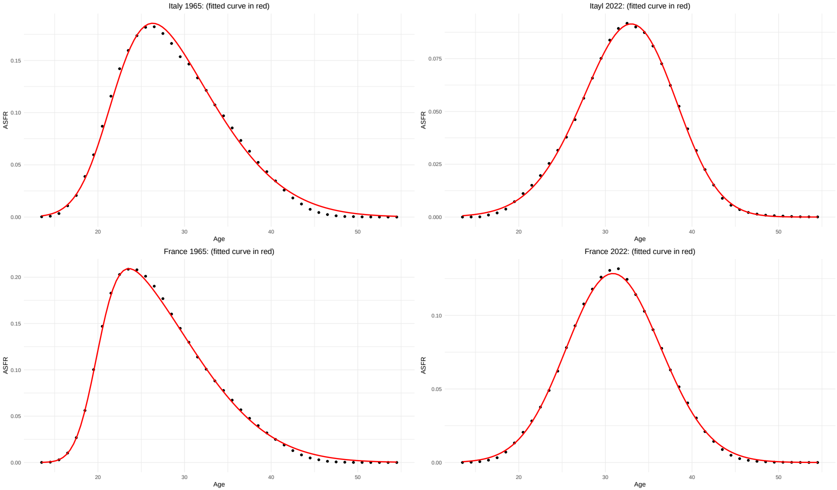


Figure: Fertility schedule of Italy and France (fitted by a SN)

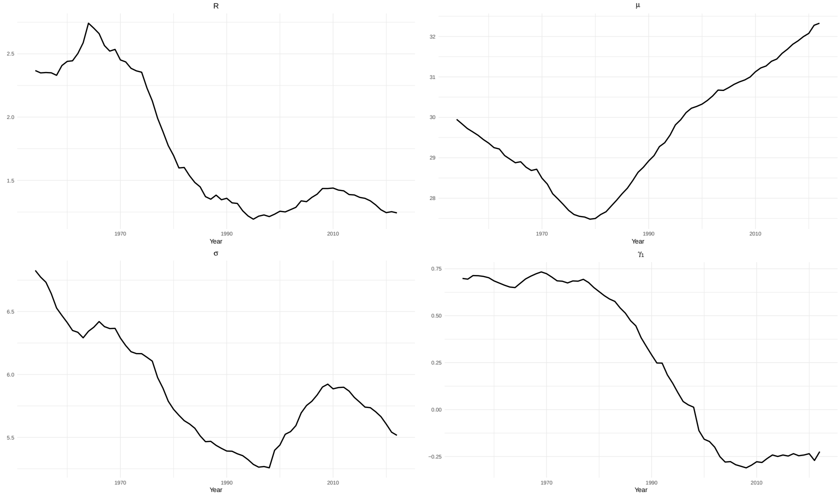


Figure: Skew normal (cp) parameters estimates on Italy fertility (1954–2022).

Tempo distortion pf Period Fertility

What is “tempo distortion”?

- Fertility is usually measured by *Period* Total Fertility Rate (TFR)

$$PTFR = (t) = \int_{\mathcal{R}^+} g(x; R, \theta_2, \ldots, \theta_r) = R \int_{\mathcal{R}^+} h(x; \theta_2, \ldots, \theta_r) dx$$

- However the integral/sum is calculated on women belonging to different cohorts, in case of fertility displacement of some of these cohorts, the PTFR can suffer from distortion even if the intensity of fertility does not change across cohorts

What is “tempo distortion”?

Università degli Studi di Padova

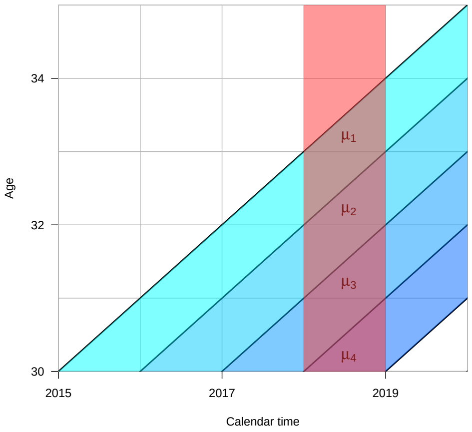


Figure: Graphical representation of tempo distortion

What is "tempo distortion"?

- Let us assume that CTFR is constant over cohorts but the timing changes and consider $h(x; \mu(\tau), \sigma(\tau), \gamma(\tau))$ as the fertility rate at age x for cohort born in τ we have that

$$CTFR(\tau) = R \int_{\mathcal{R}^+} h(x; \mu(\tau), \sigma(\tau), \gamma(\tau)) dx = R.$$

- The value of CTFR does not change over τ because $h(\cdot)$ is a probability density function, therefore its integral over the whole support is 1. However

$$PTFR(t) = R \int_{\mathcal{R}^+} h(x; \mu(t-x), \sigma(t-x), \gamma(t-x)) dx \neq R.$$

Adjusting for mean, variance and skewness variations

- we can adjust the age-specific fertility rate f_x^t getting the value that would have been observed if the *timing* of fertility schedule had remained stable:

$$\tilde{f}_x^t = \frac{f_x^t}{\lambda^{t_0}(x)},$$

- where $\lambda^{t_0}(x,)$ is the ratio between $h(x|\mu_t, \sigma_t, \gamma_t)$ and $h(x|\mu_{t_0}, \sigma_{t_0}, \gamma_{t_0})$.

Adjusting for mean, variance and skewness variations

Università degli Studi di Padova

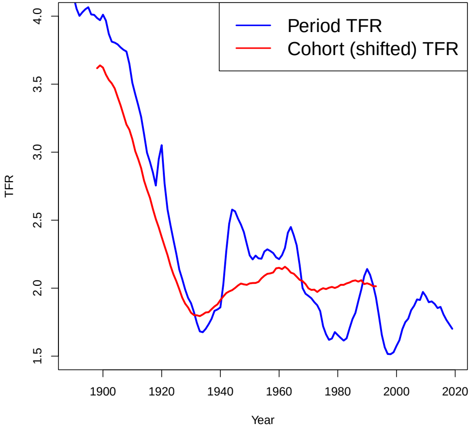


Figure: Period and (shifted) cohort fertility, Sweden, 1899–2020

Adjusting for mean, variance and skewness variations

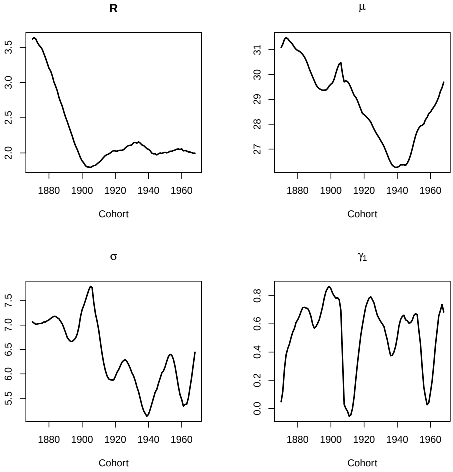


Figure: Skew normal parameters estimates, Sweden, cohorts 1872–1968

Adjusting for mean, variance and skewness variations

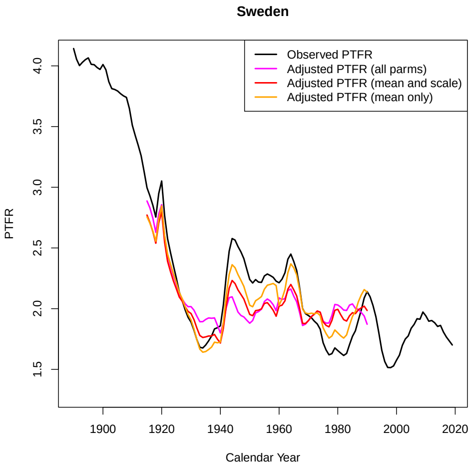


Figure: Sweden PTFR and adjusted PTFFR, years 1890–2020

Skewness of old age mortality age profiles

Lexis or Pearson?

Premature mortality

- Lexis: Premature deaths as the difference between an hypothetical "normal" death curve (if only senescence operates) and the observed, usually left-skewed, one
- For Pearson, the observed asymmetric distribution of deaths in a life table is itself the "normal" distribution
- Premature mortality is thus not a statistical "residual" but a *defined component* of mortality curve

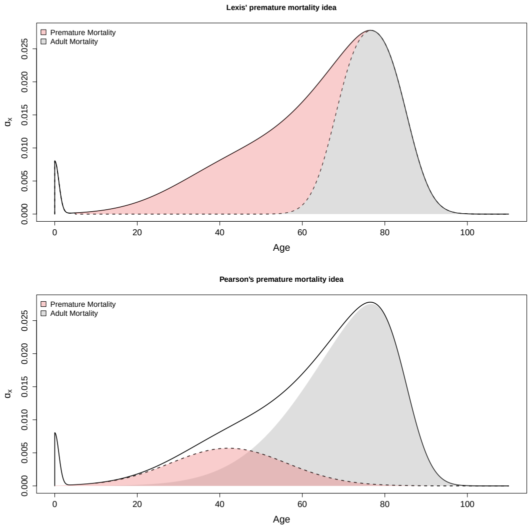


Figure: Graphical representation of Lexis and Pearson ideas of premature mortality

A Mixture model

$$f_M(x, \xi_M, \omega_M, \lambda_M) = \overbrace{\frac{2}{\omega_M}\phi\left(\frac{x-\xi_M}{\omega_M}\right)\Phi\left(\lambda_M\frac{x-\xi_M}{\omega_M}\right)}^{\text{Senescent mortality}} \tag{1}$$

$$f_m(x, \xi_m, \omega_m) = \overbrace{\frac{1}{\omega_m}\phi\left(\frac{x-\xi_m}{\omega_m}\right)}^{\text{Premature mortality}} \tag{2}$$

$$f_I(x) = \overbrace{\frac{\sqrt{2}}{\pi}\exp\left(-x^2\right)}^{\text{Infant mortality}} \tag{3}$$

$$\begin{aligned} d(x) &= \eta \cdot f_I(x) + (1-\eta)\cdot\alpha\cdot f_m(x, \xi_m, \omega_m) + \\ &+ (1-\eta)\cdot(1-\alpha)\cdot f_M(x, \xi_M, \omega_M, \lambda_M) \end{aligned}$$

The hierarchical version of the mixture model

$$d_j(x) = \alpha_j \cdot f_j^m(x, \mu_j^m, \sigma_j^m) + (1 - \alpha_j) \cdot f^M(x, \mu^M, \sigma^M, \gamma^M)$$

- The old-age component is common
- The premature component varies across regions
- Infant mortality is disregarded (only mortality for age > 5 is considered)

Preamture mortality and Covid-19

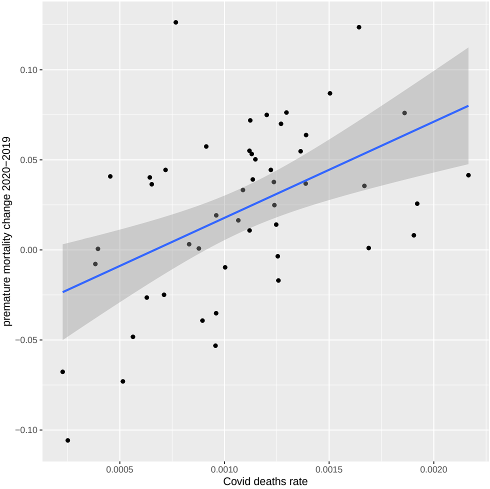


Figure: Premature mortality change (2019-2020) and Covid mortality in US states

Skewness in mortality

A matter of selection?

- One key argument of Pearson in its critique to Lexis is selection
- individuals is not a random sample of the original birth cohort; it is a highly selected subset that has survived a lifetime of risks.
- This selected group of robust elderly does not die according to a symmetrical distribution around a “normal” lifespan.

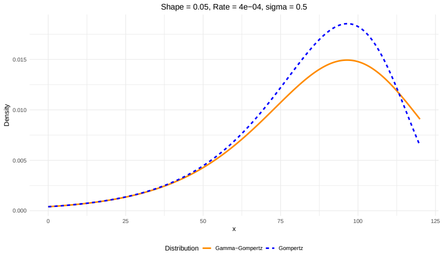
Gompertz and Gamma-Gompertz

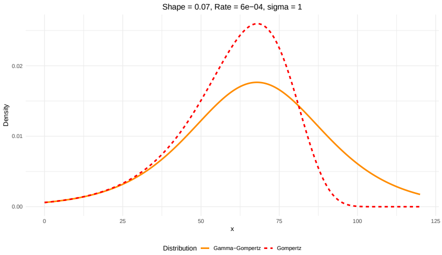
- Recall that Vaupel introduced a Gamma frailty to take into account of selection in Gompertz model

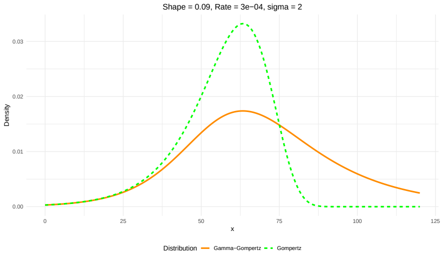
$$h(t_i|Z_i) = \alpha e^{\beta t_i} \cdot Z_i$$

- Gompertz is the "normal" distribution of deaths by age but indivduals are a mixture of differently "frail" subsample
- Thus the marginal distribution of observed population has a different shape (Gamma-Gompertz)

Gompertz and Gamma-Gompertz







Skew normal additive representation

- Recall that a stochastic representatio of SN is

$$Y = \delta|Z_1| + \sqrt{1-\delta^2}Z_2$$

  wit $Z_1, Z_2 \sim N(0,1)$
- δ is proportional to skewness parameter
- A negative skewness suggests we are dealing with a subsample with lower average length of life

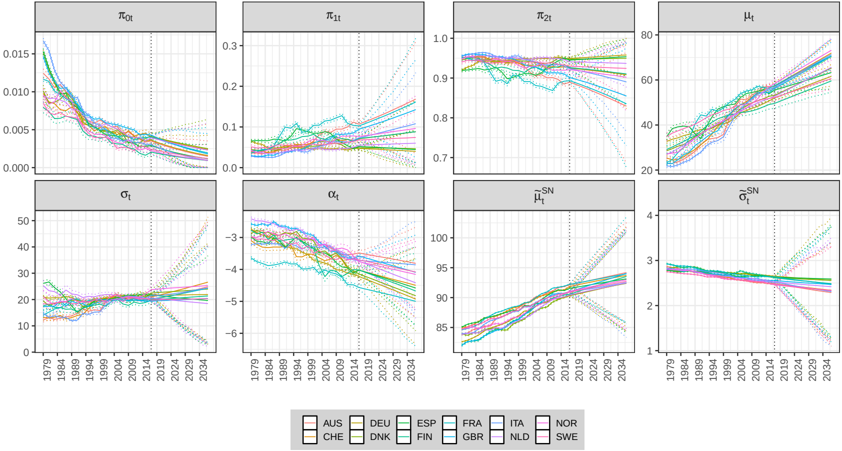


Figure: Mixture model parameters estimate and forecasts for some countries, source: Aliverti et al (2022)

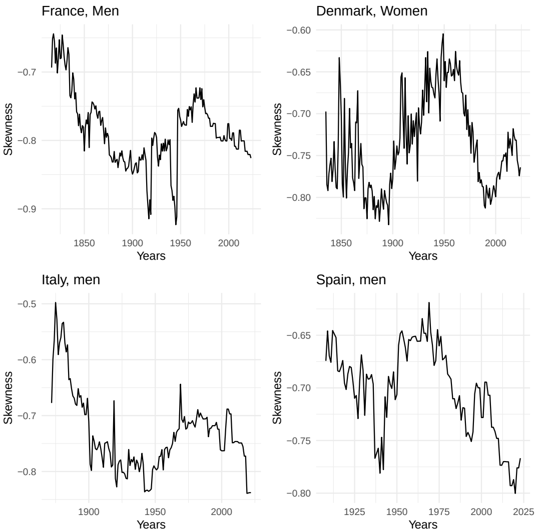


Figure: Skewness parameter of mortality schedules in four countries, source: HMD

To recap

- Skewness in (cohort) fertility: evidence of fertility displacement
- In some cases such displacement does not involve mean but shape and/or scale parameters
- Skewness in mortality: evidence of heterogenous subpopulations
- The higher the skewness the lower the selection?

References

- S, Mazzuco, Scarpa B. (2015).
  Fitting age-specific fertility rates by a flexible generalized skew normal probability density function *Journal of the Royal Statistical Society - Series A*

- S, Mazzuco, Zanotto L. (2025).
  Tempo effects in period TFR *Demographic Research*

- Aliverti E., Durante D., Scarpa S. (2020).
  Projecting Proportionate Age–Specific Fertility Rates via Bayesian Skewed Processes, in *Developments in Demographic Forecasting*, Springer

- Aliverti E., Scarpa B., and Mazzuco, S. (2022).
  Dynamic Modelling of Mortality Via Mixtures of Skewed Distribution Functions. *Journal of the Royal Statistical Society - Series A*.

- Zanotto L., Canudas-Romo V. and Mazzuco S. (2020).
  A mixture-function mortality model: illustration of the evolution of premature mortality. *European Journal of Population*.

- Mazzuco S., Suhrcke, M. and Zanotto L. (2021).
  How to measure premature mortality? A proposal combining “relative” and "absolute" approaches. *Population Health Metrics*.

Acknowledgements

- Project: CARONTE – PRIN 2022 (Prot. 2022KBTEBN)
- Funded by the EU – NextGenerationEU, Mission 4

Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

Questions welcome.