Fertility Models

POP 502 / ECO 572/ SOC 532 • SPRING 2017

We now discuss age patterns of fertility, with a brief review of the Henry, Coale, Coale-Trussell, Page, Brass and Schmertmann models. The textbook deals with this subject in Section 9.3.

Natural Fertility and Control

Henry discovered that many natural fertility populations (where there is no conscious attempt to control the number of children) differed in the level of fertility but had a similar age profile. We call this the natural fertility pattern and embody it in a schedule $n(a)$. The actual age-specific rates in a natural fertility population are proportional to this schedule. The shape of the schedule is represented by the top line in Figure 1.

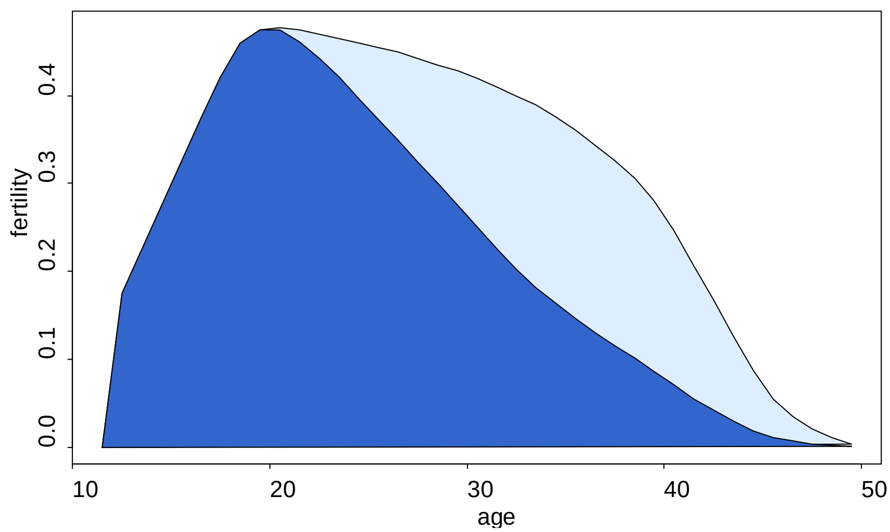

Figure 1. Natural and Marital Fertility

Coale discovered that populations that controlled their fertility (consciously limiting the number of children) exhibit a typical pattern of deviation from natural fertility, with departures in the log scale increasing with age according to a typical schedule $v(a)$. This led him to propose a model where marital fertility at age a is

$$m(a) = M\, n(a) e^{-mv(a)}$$

Here M is a parameter representing the level of marital fertility, and m is a parameter representing the degree of control. Note that dividing by the natural fertility schedule $n(a)$ and taking logs we obtain

$$\log \frac{m(a)}{n(a)} = \alpha + \beta v(a)$$

with intercept $\alpha = \log(M)$ and slope $\beta = -m$. This makes it very easy to check visually if a given fertility schedule follows Coale's model. The model is also easy to fit by simple linear regression. An even better approach is to use Poisson regression, as suggested by Bröstrom and Trussell.

The lower line in Figure 1 shows a marital fertility schedule where $M = 1$ and $m = 1$. The lighter area shows the extent to which fertility falls below natural fertility as a result of control. The textbook shows a simple application to data from Mali in 1955-6 with $M = 0.76$ indicating a level of natural fertility 24% below Henry's level, and $m = 0.189$ showing little evidence of control. The fit is not particularly good.

General Fertility by Age

Coale and Trussell combined Coale's model of marital fertility with the Coale-McNeil model of marriage, writing the general fertility rate at age a as the product of the proportion married at that age by the age-specific marital fertility rate, assuming no fertility outside marriage. (This assumption is a lot more reasonable if we define "marriage" to include both legal and consensual unions.) Thus

$$f(a) = cG_0\left(\frac{a-\mu}{\sigma}\right)Mn(a)e^{-m\,v(a)}$$

It is easy to see that the parameters c and M are not separately identified, so we can't distinguish the level of natural fertility from the proportion who eventually marry from age-specific fertility rates alone. (Unless, of course, we have data on proportions married and marital fertility and fit the two components of the model separately.)

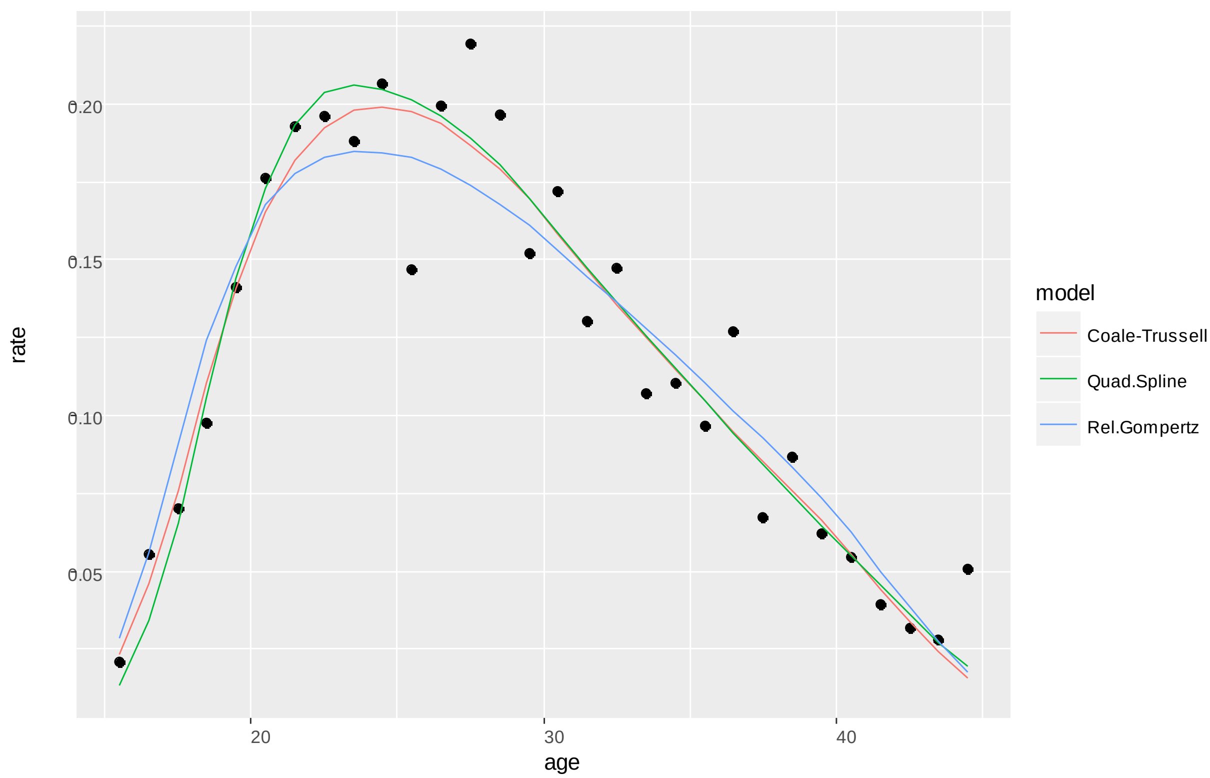

Figure 2. Age-specific Fertility Rates and Model Fits, Brazil DHS 1986

In a paper with Philipov we show how to fit this model using maximum quasi-likelihood procedures. A simple alternative is to use non-linear least squares. Figure 2 shows our fit to data from the Brazil Demographic and health Survey of 1986 using births in the four years before the survey. The observed rates are noisy but the model does an excellent job smoothing the data. (The figure also shows two other models discussed below.)

The estimated parameters (μ, σ, M, m) are (20.52, 5.10, 0.591, 0.771), indicating that average age of entry into union (exposure) is 20.52 with a standard deviation of 5.10, and that there is substantial fertility control among those in union (exposed). Because the Brazilian DHS includes information about unions we were able to fit the two components of the model separately. We got (22.12, 5.39, 0.954) from the union component and (0.738, 0.993) from the marital fertility component. The separate models imply later entry into risk and more control, but do not fit the observed rates as well as the combined model (details not shown).

Marital Fertility by Age and Duration

Page proposed a model of marital fertility where the level of natural fertility depends on a woman's age as in the previous models, but the degree of control depends on the duration of marriage instead of age. She further discovered that the pattern of departure from natural fertility viewed as a function of union duration rather than age was linear in the log scale, so no special schedule of control was needed. This leads to the following model of marital fertility by age and duration

$$m(a,d) = Mn(a)e^{\beta d}$$

Or, dividing by the natural fertility schedule and taking logs.

$$\log \frac{m(a,d)}{n(a)} = \alpha + \beta d$$

where $\alpha = \log(M)$. In work with Cleland we fitted this model to data from the WFS and showed that the degree of control parameter β is strongly correlated with contraceptive use, particularly use for limiting, and the level of natural fertility parameter α is correlated with breastfeeding duration as well as contraceptive use for spacing. We therefore called α and β the limiting and spacing parameters. In subsequent work we used this model to study social determinants of fertility in terms of their effects on limiting and spacing behavior by letting each of the parameters depend on covariates.

The Relational Gompertz Model

Brass proposed a relational Gompertz model of fertility. The basic idea is to transform the proportion of cumulative fertility achieved by age a using a log-log transformation

$$Y(a) = -\log(-\log\left(\frac{F(a)}{F}\right))$$

where F is short for $F(\infty)$, representing the total fertility rate, and then assume that the transformed schedule is a linear function of a standard schedule

$$Y(a) = \alpha + \beta Y_s(a)$$

which was derived by Booth and is available in the *Tools for Demographic Estimation* website. (Interestingly the standard was derived by examining a large collection of Coale-Trussell schedules.)

The transformation in the first equation above is called a *gompit* and is related to the Gompertz growth function, just like a *logit i*s related to the logistic growth function. The Gompertz growth function in turn, is closely related to the Gompertz survival function.

The model is very easy to fit by OLS if we accumulate the observed rates and then divide by the TFR to obtain proportion of total fertility achieved at each age, and has applications in indirect estimation. Unfortunately it doesn't fit the Brazilian data very well, as you can see in Figure 2.

The Quadratic Spline Model

Schmertmann (2003) proposed a quadratic spline model with four "graphically intuitive" parameters, a, P, H and R. The first three represent the ages at which fertility rises above zero, reaches its peak, and falls to half its peak level, and the last one is the peak fertility.

Figure 2 shows that the model fits the Brazilian data reasonably well. (The residual sums of squares are 0.0087 for Coale-Trussell, 0.0090 for Schmertmann and 0.0110 for Brass, but note that the relational Gompertz model has three parameters and the other two models have four each.) The quadratic spline index ages are 13.85, 23.18 and 35.66, and the peak level is 0.206.

The model is a quadratic spline, so it can be written as

$$f(a) = R \sum_{k=0}^{4} \theta_k (a - \xi_k)_+^2$$

where ξ_k are the knots and θ_k the coefficients for $k = 0, \ldots, 4$. These parameters can be obtained from the index ages by solving a system of linear equations which imposes a number of constraints to reduce the number of shape parameters from 10 to three.

A nice feature of the model is that it can be integrated analytically to obtain cumulative fertility as a cubic spline

$$F(a) = \frac{R}{3} \sum_{k=0}^{4} \theta_k (a - \xi_k)_+^3$$

When working with grouped data I recommend computing the rate for ages $(x, x+n)$ as the difference $F(x+n) - F(a)$, which is more accurate than the mid-point rate $f(x + \frac{n}{2})$.