Probabilistic projection of UK kinship

Joe William Barrass Butterick[1], Joanne Ellison[1], Peter W. F. Smith[1], Erengul Dodd[1], Jakub Bijak[1], Jonathan J. Forster[2], and Jason Hilton[1]

[1]University of Southampton
[2]University of Warwick

1 Theoretical emphasis

By the year 2070, the number of people aged 80 and above in the UK is expected to double (ONS, 2024). Family is at present the main source of late-life care, however, as changes in mortality and fertility consistently alter the inter-generational structures of family, so does the nature of kin available to care.

The ages at which people die and reproduce are intertwined with a continually changing world. Increasing living standards are prolonging life-expectancy. Improved access to labour market opportunities are delaying mean age of childbearing. These socio-demographic factors are re-shaping kinship networks. For example, an emergent "double sandwich" generation (so-called due to 4 generations co-existing at one time) of 40-50 year-old individuals must simultaneously care for grand-children – thus enabling their children to work, and parents – who on average being aged 80 plus, require support (Butterick et al., 2024). It is of great policy interest to understand in 50 years time which kin are available to offer an individual financial, emotional, and other forms of support.

One way to estimate kinship networks is through the matrix population model proposed by Caswell (2019), a theoretical simple and computational efficient framework. The matrix model takes as input demographic rate data, and gives as output expected age-distributions of kin, relative to a reference individual; "Focal". Since its inception, the theoretical foundations of matrix-based kinship have been extended to account for time-varying demographic rates (Caswell & Song, 2021), multi-state (Caswell, 2020) and two-sex populations (Caswell, 2022), and more recently stochastic population dynamics (Caswell, 2024). With regards to applications, recent work in Alburez-Gutierrez et al. (2023) produced probabilistic global kinship forecasts.

The goal of this research is to probabilistically project UK kinship networks into the future by combining the models developed in Caswell & Song (2021) and Caswell (2022). We inform our proposed model with historic (1938-2020) demographic rates taken from the Human Fertility Collection (HFC) and Human Mortality Database (HMD). Projected (2021-2070) demographic rates are sourced from Bayesian statistical models. Rather than following Alburez-Gutierrez et al. (2023) who apply an "androgynous approximation" (using female fertility for both sexes), we apply sex-specific mortality and fertility rates to provide more accurate inferences of kin. Uncertainty in our forecast kin estimates naturally emerges through the uncertainty in the statistical forecasts.

In what follows, we demonstrate how kinship changes both over time and age of Focal. We additionally demarcate extremal upper and lower bounds for the numbers of kin in the future. Our results will have profound importance for policy planning.

2 Methods

Here we construct kinship using the models proposed in Caswell & Song (2021) and Caswell (2022). Consider a population defined at time t and composed of sexes $i = \{\mathrm{f}, \mathrm{m}\}$. Let $\mathbf{F}_t^i$ be a matrix comprising age-specific-fertility-rates on the top row. Let $\mathbf{U}_t^i$ be a matrix with a sub-diagonal of age-specific-survival-probabilities and zeros elsewhere. Suppose Focal is of age x at time t, and let $\mathbf{k}_t^i(x)$ represent an age-distributed vector of the expected numbers of some particular $i-$sex kin (e.g., the j^{th} entry of $\mathbf{k}_t(x)^{\mathrm{m}}$ is a scalar representing the expected number of age-class j male kin of Focal of age x at time t). We project a block-structured $\tilde{\mathbf{k}}_t(x) = (\mathbf{k}_t^{\mathrm{f}}(x) \mid \mathbf{k}_t^{\mathrm{m}}(x))^{\dagger}$ through

$$\tilde{\mathbf{k}}_{t+1}(x+1) = \begin{cases} \tilde{\mathbf{U}}_t\tilde{\mathbf{k}}_t(x), & \text{(no recruitment)} \\ \tilde{\mathbf{U}}_t\tilde{\mathbf{k}}_t(x) + \tilde{\mathbf{F}}_t^*\tilde{\mathbf{k}}_t^{(\mathrm{subs})}(x) & \text{(dependent recruitment)} \\ \tilde{\mathbf{U}}_t\tilde{\mathbf{k}}_t(x) + \tilde{\mathbf{F}}_t\tilde{\mathbf{k}}_t^{(\mathrm{subs})}(x) & \text{(independent recruitment)} \end{cases} \tag{1}$$

where the matrices

$$\tilde{\mathbf{U}}_t = \mathbf{U}_t^{\mathrm{f}} \oplus \mathbf{U}_t^{\mathrm{m}}, \;\; \tilde{\mathbf{F}}_t = \begin{pmatrix} (1-\alpha)\mathbf{F}_t^{\mathrm{f}} & (1-\alpha)\mathbf{F}_t^{\mathrm{m}} \\ \alpha\mathbf{F}_t^{\mathrm{f}} & \alpha\mathbf{F}_t^{\mathrm{m}} \end{pmatrix}, \;\; \tilde{\mathbf{F}}_t^* = \begin{pmatrix} (1-\alpha)\mathbf{F}_t^{\mathrm{f}} & 0 \\ \alpha\mathbf{F}_t^{\mathrm{f}} & 0 \end{pmatrix} \tag{2}$$

respectively account for survival of females and males independent of one another; reproduction of both sexes; and reproduction of females only, while α reflects the proportion of male newborns. For a detailed description of the kin projection methods see Caswell (2022). Below we recapitulate the salient features of the framework.

At birth ($x = 0$), Focal will experience initial (possibly zero) age-distributions for each kin within its network. Thereafter ($x \geq 1$), consider Focal advances age-class $x \to x+1$. Kin of Focal in age-class $x+1$ are the sum of firstly: kin of Focal in age-class x which survive, as defined through the prevailing population mortality rates, and secondly: kin born between Focal moving from age-class x to $x+1$, to a "subsidiser" kin $\tilde{\mathbf{k}}^{(\mathrm{subs})}$, as defined through the prevailing population fertility rates.

With the above-mentioned conceptual groundwork, the RHS of Eq (1), from the top line to bottom describes the time-change in kin under three distinct cases: (i) kin which Focal cannot acquire any more of during its life (e.g., older siblings); (ii) kin which Focal can acquire during its life, but which cannot be procured independently by female and male subsidisers (e.g., younger siblings with parents); (iii) kin which Focal can acquire during its life, which are independently procured by both female and male subsidisers (e.g., nieces and nephews). Notice that case (ii) assumes a female dominant population, only counting descendants of the female direct ancestors of Focal. This assumption circumvents the possibility that Focal's parents or grand-parents reproduce independently.

The above-described model applies 1-year age and time intervals. Historic age-specific-fertility-rates (ASFR) are taken from the HFC (Human Fertility Collection, 2024), while historic age-specific-survival-probability (ASSP) are sourced from the HMD (Human Mortality Database, 2024), both spanning a time period $t = 1938 - 2020$. In Section (3), we assume that future rates cover the years $t = 2021 - 2070$, with ASSP forecast under the Bayesian model proposed by Hilton et al. (2018), while fertility projections are obtained using a Lee-Carter type model, fitted to historical data using singular value decomposition and forecast using an Auto-Regressive-Moving-Average model (Lee, 1993). In Section (4) we implement a more sophisticated projection model, estimating female parity-specific ASFR based on the model of Ellison et al. (2023) and mapping to male.

In both moelds, median point estimates and uncertainty intervals for kin are produced by repeatedly informing the matrix model with independent time-trajectories for ASFRs and ASSPs, as sampled from the statistical models.

3 Two-sex forecasts

We present preliminary model outputs for UK kinship over historic timescale (1938-2020) and forecast timescale (2021-2070) using 500 independent time-trajectories of forecast ASFRs and ASSPs. The model output is amenable to analysis by period, time and cohort, as detailed below.

Figure 1 gives an example of the accumulated number of siblings for different ages across the life course of Focal, over different periods in time. Note the widening uncertainty bound at the early ages of Focal as we move forwards in time period. This effect has two sources: first, a widening of uncertainty in the ASFR forecast, and second, more years in which Focal's mother's reproduction is defined through uncertainty in fertility rate. For instance, in 2030, a typical 12 year-old Focal (born in 2018) experiences uncertainty in its number of younger siblings as consequence of its mother producing these kin under uncertain fertility rates; only between 2021-2029. Whereas a 12 year-old Focal in 2050 (born in 2032) experiences greater uncertainty in younger siblings since its mother produces these kin both (i) under wider uncertain intervals of fertility and (ii) for a longer duration (right throughout Focal's life).

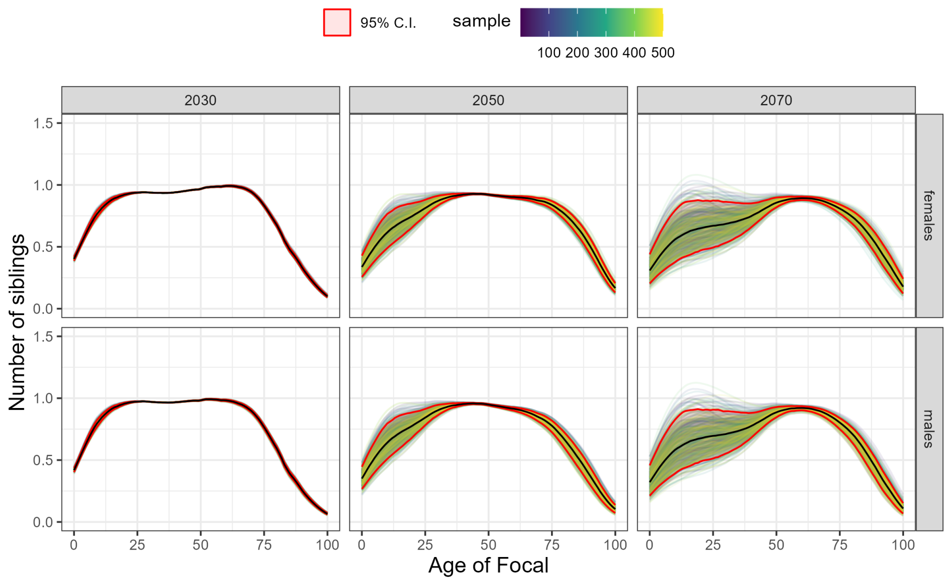

Figure 1: Median (black) and 95% CI (red) number of siblings as a function of age of a typical Focal's, compared across different periods of time. Uncertainty bounds calculated using 500 samples of the statistical forecasts (coloured).

The median expected number of kin a typical Focal individual will acquire as a function of their year of birth is shown in Figure 2. We see monotonic decreases in the expected number of offspring and grand-offspring as we move to more recent cohorts. This is explained by an overall lowering (of both observed and predicted) total fertility from 1980 to 2050. Contrastingly, considering parents, as we move forwards by cohort we observe monotonic increases in the probability these kin are alive at any age over the life course of Focal, as explained by improvements in mortality schedules over time.

The accumulated kin that Focal – of fixed age – will experience as a function of time are illustrated in Figure 3. Here, Focal is assumed 50 years old. We see that all kin-types exhibit a qualitatively similar temporal trajectory. Assuming that Focal's aunts and uncles are around age 80, it makes sense that uncertainty in the expected number of these kin manifests itself after 2021; simply through variability in mortality rate (these kin are born under historic fertility rates). The overall number of aunts and uncles peaks around 2040; we assume they are born in the 1960s under the highest fertility rates of the 20th century. Focal's cousins, as subsidised by aunts and uncles, follows a similar pattern but with propagation of numbers and delayed peak. We assume cousins are of similar age to Focal (around 50) and thus should be born on average under historic (and deterministic) fertility predictions. Uncertainty in the number of these kin comes primarily from mortality rate forecasts. Contrastingly, nieces and nephews of Focal can be born under uncertain fertility rates: assuming Focal's sisters and brothers have children around age 30, nieces and nephews will be around age 20. By the year 2041 we expect nieces and nephews to be born under variable fertility forecasts, and therefore exhibit uncertainty in their numbers.

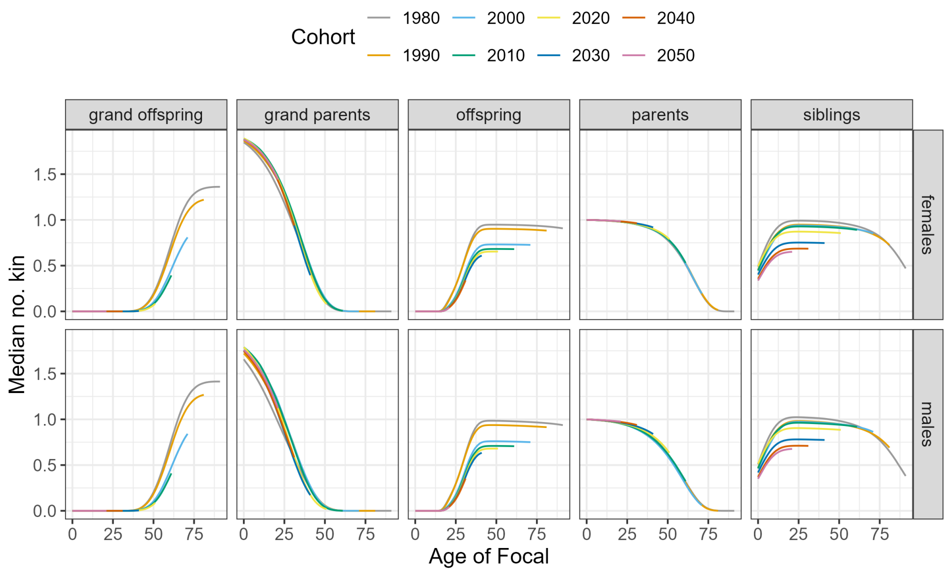

Figure 2: Median number of kin as a function of Focal's cohort of birth.

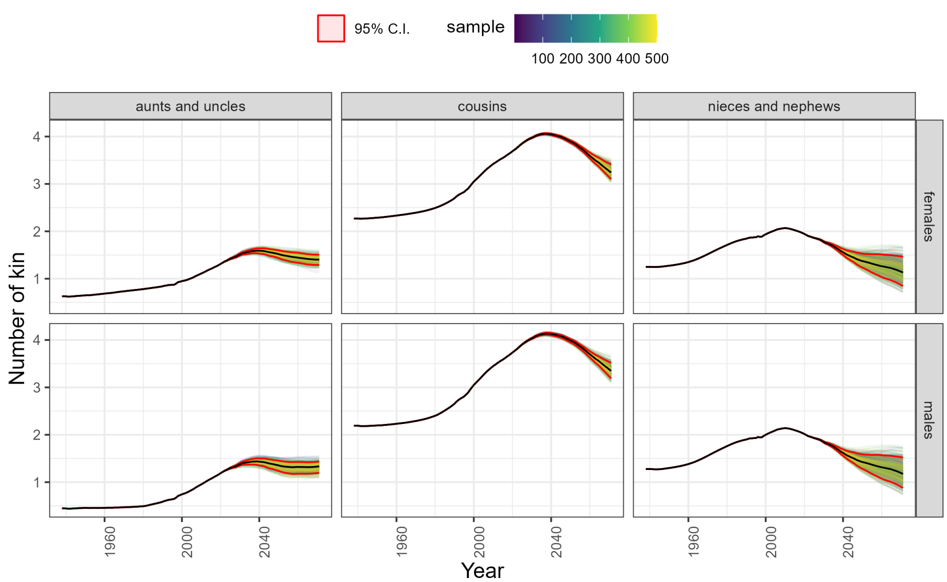

Figure 3: Median (black) and 95% CI (red) accumulated number of kin a typical 50 year-old Focal will experience, as a function of time. Uncertainty intervals produced using 500 samples of the statistical forecasts (coloured)

4 Parity-specific forecasts

Here, we present model output broken by sex and parity. We apply fully Bayesian forecasts for ASFRs and ASSPs over timescale (2023-2070). Fertility forecasts are sex and parity specific. Mortality forecasts are sex specific; we do not assume parity affects the probability of survival. At present, due to availability of parity specific birth data, we run the kinship model on a historic timescale (2000-2022). Ongoing work is estimating parity specific fertility over the years 1938-1999. Once this estimation is complete, the kinship model will be run over a longer timescale providing for a more realistic unfolding of kin-networks. The model output is amenable to analysis by sex, parity, period, time, and cohort, as detailed below.

Figure 4 gives an example of the parity of (a female) Focal, by age of Focal assuming that Focal is born in 2010. Note that because we forecast to 2070, a 2010 cohort individual is projected to age 60. By the age of 60, we expect Focal to be more likely in parity 2 than any other. We see that Focal is in parity 0 until she reaches reproductive age (15 years old), at which point we have non-zero probabilities of her transitioning to parity 1. By the age of 17 Focal has a non-zero probability of moving to parity 2, and by 18 a non-zero probability of moving to parity 3+. We see widening uncertainty in the parity of Focal caused by rate uncertainty, up to age 55, at which point her reproductive life is over.

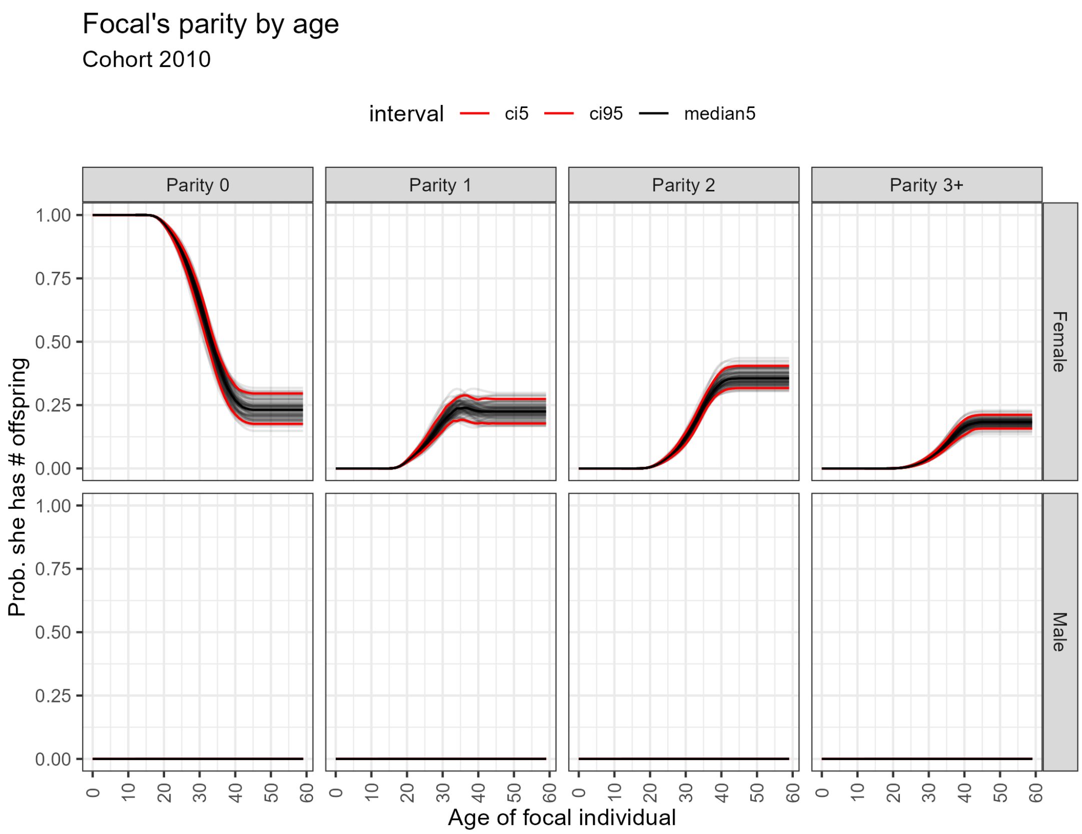

Figure 4: Median (black) and 95% CI (red) parity of Focal as a function of age of a typical Focal born in 2010. Uncertainty bounds calculated using 500 samples of the statistical forecasts (grey).

Figure 5 gives an example of the accumulated number, and parity, of offspring of Focal, by age of Focal assuming that Focal is born in 2010. Note that because we forecast to 2070, a 2010 cohort individual is projected to age 60. By the age of 60, we expect most of her offspring to be in parity 0. We do not expect Focal to not have offspring in parity 1 until she is around 40 years old, and we do not expect Focal to have offspring in parity 2 until she is around 45. Again, we see widening uncertainty in the number of (parity 0) offspring as Focal gets older: both her accumulated fertility increases and the rate uncertainty increases.

We could also explore other kin-relations, for example in Figure 6 we show the accumulated number, and parity, of younger sisters of Focal, by age of Focal assuming that Focal is born in 2010.

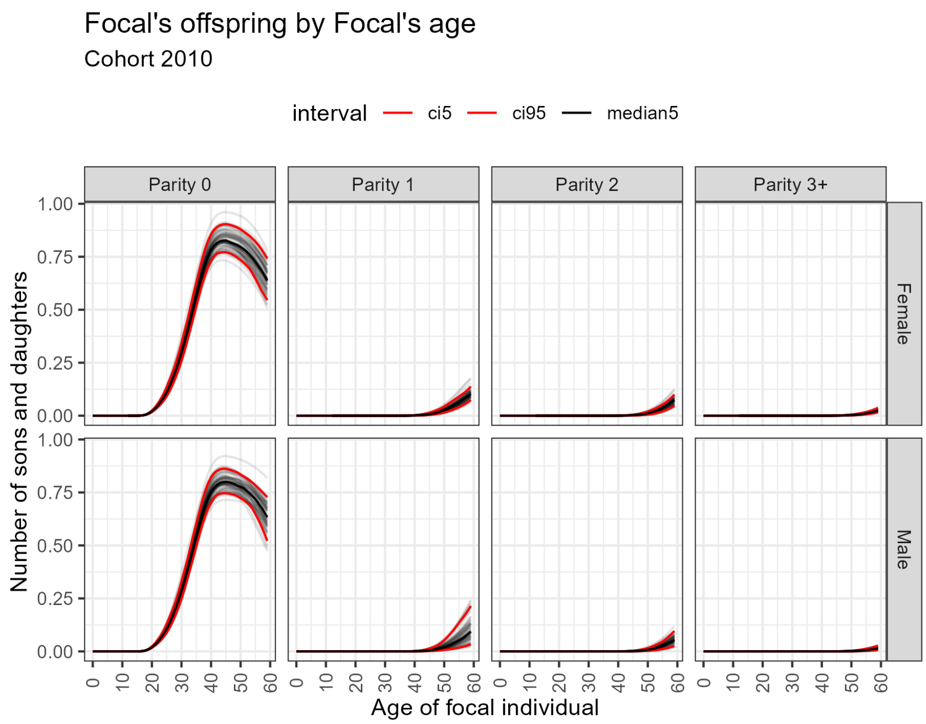

Figure 5: Median (black) and 95% CI (red) number and parity of offspring as a function of age of a typical Focal born in 2010. Uncertainty bounds calculated using 500 samples of the statistical forecasts (grey).

5 Further investigation

Further work for this paper will involve a validation of the model using data from the English Longitudinal Survey of Ageing (Banks et al., 2024).

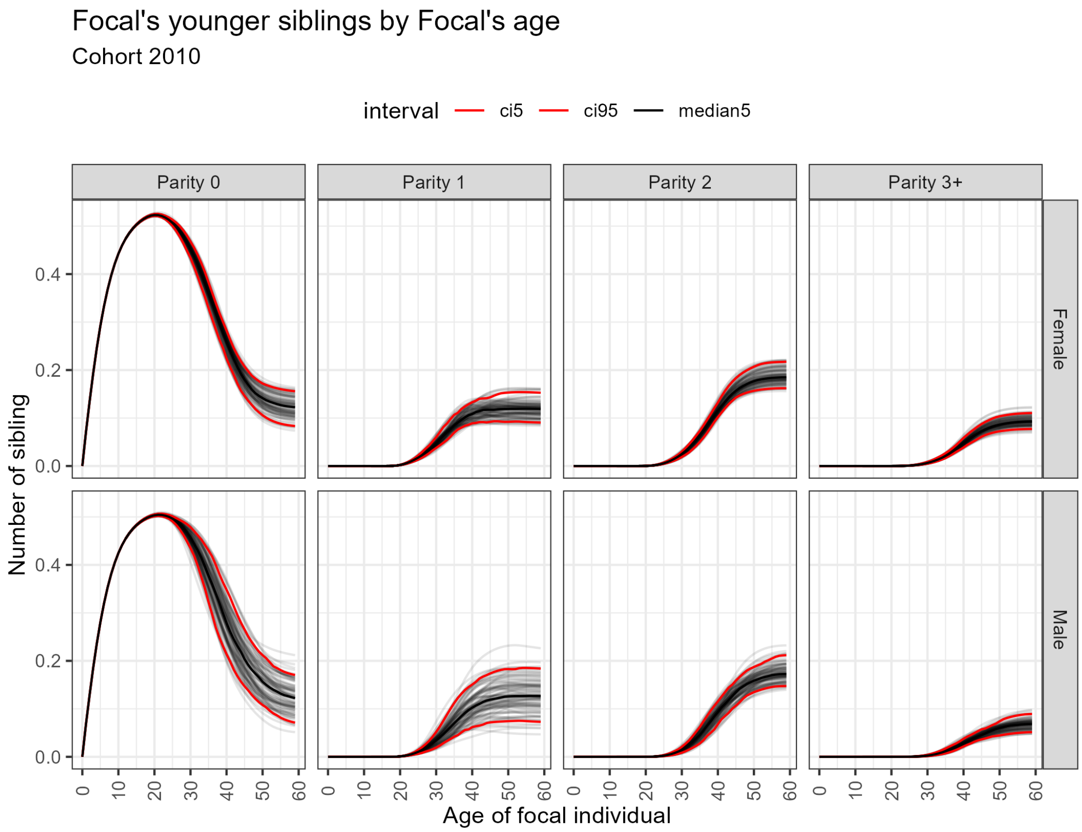

Figure 6: Median (black) and 95% CI (red) number and parity of Focal's younger sisters as a function of age of a typical Focal born in 2010. Uncertainty bounds calculated using 500 samples of the statistical forecasts (grey).

References

Alburez-Gutierrez, D., Williams, I., & Caswell, H. (2023). Projections of human kinship for all countries. *Proc. Natl. Acad. Sci.*, *120*(52).

Banks, J., Batty, G., David, B. J., Coughlin, K., Crawford, R., Marmot, M., Nazroo, J., Oldfield, Z., Steel, N., Steptoe, A., Wood, M., & Zaninotto, P. (2024). English Longitudinal Study of Ageing: Waves 0-10, 1998-2023.
URL https://beta.ukdataservice.ac.uk/datacatalogue/

Butterick, J. B., Evandrou, M., Falkingham, J., & Hilton, J. (2024). Changes in kinship: implications for the availability of kin to care. Project report.

Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Res.*, *41*, 679–712.

Caswell, H. (2020). The formal demography of kinship II: Multistate models, parity, and sibship. *Demographic Res.*, *42*(38).

Caswell, H. (2022). The formal demography of kinship IV: Two-sex models and their approximations. *Demographic Res.*, *47*.

Caswell, H. (2024). The Formal Demography of Kinship V: Demographic stochasticity, variance, and covariance in the kinship network. *BioRxiv [preprint]*.

Caswell, H., & Song, X. (2021). The formal demography of kinship III: Kinship dynamics with time-varying demographic rates. *Demographic Res.*, *45*.

Ellison, J., Berrington, A., Dodd, E., & Forster, J. J. (2023). Combining individual- and population-level data to develop a bayesian parity-specific fertility projection model. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *73*(2), 275–297.
URL https://doi.org/10.1093/jrsssc/qlad095

Hilton, J., Dodd, E., Forster, J. J., & Smith, P. W. F. (2018). Projecting UK Mortality by Using Bayesian Generalized Additive Models. *J. R. Stat. Soc., C: Appl. Stat.*, *68*(1), 29–49.

Human Fertility Collection, . (2024). Max Planck Institute for Demographic Research (Germany) and Vienna Institue of Demography (Austria).
URL https://www.fertilitydata.org.

Human Mortality Database, . (2024). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).
URL https://www.mortality.org.

Lee, R. D. (1993). Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level. *Int. J. Forecast.*, *9*(2).

ONS (2024). National Population Projections.
URL https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/datasets/tablea21principalprojectionukpopulationinagegroups