

Ecological harshness mediates reproductive trade-offs in a great tit population

Louis Bliard ¹, Jordan S. Martin ², Dylan Z. Childs ³, Ella F. Cole ⁴, Ben C. Sheldon ⁴, Maria Paniw ^{1,5},
Arpat Ozgul ¹

¹ Department of Evolutionary Biology and Environmental Studies, Zurich University, Zurich,
Switzerland

² Department Fish Ecology and Evolution, Swiss Federal Institute of Aquatic Science and
Technology (EAWAG), Kastanienbaum, Switzerland

³ Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

⁴ Edward Grey Institute, Department of Biology, University of Oxford, Oxford, UK

⁵ Department of Conservation Biology, Estación Biológica de Doñana (EBD-CSIC), Seville, Spain

Corresponding author: bliard.louis@gmail.com

23

Abstract

24 Lack's seminal work on bird clutch sizes has spurred expansive research on reproductive trade-
25 offs, especially focusing on offspring quantity–quality trade-offs and the potential fitness
26 consequences for the parents. The environment is a critical driver of the expression of individual
27 reproductive traits, influencing them through plastic responses. However, the plasticity of
28 reproductive trade-offs themselves across environments has seldom been studied, and these
29 studies were often limited to experimental approaches and dichotomous environments. Using
30 58 years of detailed data from a great tit population, we employ the recently developed
31 'covariance reaction norm' (CRN) model to explore how continuous environmental variation
32 influences the shape of reproductive trade-offs among individuals. Our analysis reveals that the
33 correlation potentially indicative of the offspring quantity–quality trade-off is predominantly
34 stable across years, with minimal variation linked to ecological harshness during the breeding
35 season. However, the CRN also demonstrated that the correlation between offspring mass and
36 future offspring recruitment was positive, but only under harsh environmental conditions,
37 suggesting that producing larger offspring provides fitness benefits when breeding conditions are
38 suboptimal, which may reflect the importance of size for early-life competition. Altogether, this
39 work highlights that there is temporal variation in some of the phenotypic correlations, mostly
40 driven by environmental conditions, which shape the expression of offspring investment across
41 breeding seasons. Our study shows the benefits of exploring old ecological questions in the light
42 of new statistical methods, highlighting the importance of understanding how environmental
43 variation shapes the expression of life history trade-offs and the evolution of plasticity in
44 reproductive strategies.

45

46

Introduction

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

The offspring quantity–quality trade-off has been a central area of interest in life-history biology, with numerous studies exploring this long-standing topic, both theoretically (Williams, 1966; Smith & Fretwell, 1974; Fischer *et al.*, 2011), as well as empirically in various species (Lack, 1954; Sinervo & Licht, 1991; Eium & Fleming, 2000) including humans (Lawson & Mulder, 2016). All else being equal, selection should favor females producing more offspring, but these females also require enough resources to allocate to each offspring to ensure their survival, as larger offspring usually have better survival prospects (Sogard, 1997; Eium & Fleming, 2000; Marshall *et al.*, 2003; Fischer *et al.*, 2011). It was, therefore, initially proposed that each female has an inherent optimal brood size, such that it yields the highest possible recruitment, whereby any deviations from this brood size would lead to suboptimal offspring quality and thus lower overall recruitment (Lack, 1954). Overall, while some evidence agrees with Lack’s initial hypothesis (Smith & Fretwell, 1974), with manipulated increases or decreases in clutch size leading to changes in offspring condition and suboptimal recruitment as a consequence (Pettifor *et al.*, 1988, 2001), most studies found that any increase in clutch size does not necessarily decrease offspring quality, and can even lead to higher recruitment (Perrins & Moss, 1975; Boyce & Perrins, 1987). Recent work has also shown how selection on offspring quality across breeding seasons can promote rather than constrain variability in females’ clutch size in response to seasonal fluctuations (Martin *et al.*, 2025).

65

66

Several hypotheses have been put forward to explain such results, but the role of environmental variability in impacting the relationship between brood size and recruitment has

not often been addressed. One of a few attempts was made by Högstedt (1980), showing that optimal clutch size in magpies *Pica pica* is mostly mediated through variation in territory quality, with the average clutch size only being optimal for individuals occupying territories of average quality. However, models on optimal brood size have still too often ignored the role played by the environment in mediating such reproductive trade-offs (Smith & Fretwell, 1974), as in many species, individuals adjust their offspring number and size via phenotypic plasticity (Fischer *et al.*, 2011).

The effect of environmental variation over space and time on the expression of trade-offs can favour phenotypic plasticity (Service & Rose, 1985; Björklund, 2004). These plastic changes in trait expression as a response to environmental variability can lead to changes in the variance and covariance between traits (de Jong, 1989; Martin, 2025), with the latter being determined by the relative amount of variance in resource acquisition vs. resource allocation between the given traits (van Noordwijk & de Jong, 1986), with negative correlations usually expected when the variance in resource allocation is greater than the variance in acquisition. The environment influences the evolution of traits such as optimal offspring size, optimal clutch size, and their plasticity (Parker & Begon, 1986), which have since been widely studied (Braby, 1994; Fox & Czesak, 2000; Taborsky *et al.*, 2007; Allen *et al.*, 2008; Marshall *et al.*, 2008; Leips *et al.*, 2009). However, little is known about how the environment, which influences reproductive traits, can consequently lead to changes in the correlations between traits in wild populations, which could themselves be indicative of changes in the expression of trade-offs. To date, most explorations have been limited to experimental approaches with discrete environments, or a dichotomization of the underlying continuous environmental variation (Messina & Slade, 1999; Czesak & Fox,

2003; Houslay *et al.*, 2018; Mitchell & Houslay, 2021). Discretizing continuous processes is rarely justifiable (MacCallum *et al.*, 2002; Beltran & Tarwater, 2024). Historically this was done largely for practical reasons when studying trade-offs, as no modeling approach was available to easily study the influence of continuous, multivariate environmental variation on traits covariances with typical study designs in the field (Martin, 2025). However, it is likely that reproductive trade-offs in wild populations are influenced by the continuous environmental variation experienced by reproductive females. We therefore aim to use a new multivariate statistical approach that allows us to explore the role of continuous variation in environmental harshness on traits correlations, which are potentially representing reproductive trade-offs.

Here, we used one of the largest individual-based dataset of wild birds with the aim of applying a new statistical tool — the covariance reaction norm (CRN) model (Martin, 2025; Bliard *et al.*, 2025) — to revisit longstanding questions related to reproductive trade-offs faced by reproductive individuals in wild populations. The CRN model provides a novel framework to examine how continuous environmental variation influences the covariance between traits, allowing us to test longstanding hypotheses about trade-offs in a more ecologically realistic context (Martin, 2025; Bliard *et al.*, 2025). This model has been previously applied on other taxa (Soay sheep and yellow-bellied marmots) as a mostly methodological proof of concept study highlighting that context dependence in trait correlations indicative of trade-offs can be detected. This allows us to build upon this previous proof of concept study, now making use of a much larger dataset, to get new biological insights into reproductive trade-offs in the great tit population that initially formed the basis for the formulation of Lack's principle. Here, we explored the dependence on key environmental variables of the phenotypic correlations (i.e.,

among-mother) between three key traits: brood size, offspring mass, and offspring recruitment, with each trait viewed as a maternal character (i.e., analysed at the level of the brood). Overall, based on prior work, we hypothesized that we would find a negative correlation between brood size and offspring mass (Nur, 1984; Smith *et al.*, 1989), and that this correlation would become more negative in years of harsh ecological conditions (i.e., high density, low precipitation, low temperature, low beech mast), as trade-offs have often been shown to be stronger in poor conditions (Cohen *et al.*, 2020). We also hypothesized that the phenotypic correlation between offspring mass and recruitment would be positive overall, as females producing larger offspring should have an improved recruitment, with more of these offspring surviving and reproducing in following years (Perrins & Moss, 1975; Nur, 1984). Importantly, we expected this correlation to be more positive under harsh ecological conditions, as in such conditions producing larger offspring might be key for their survival. When ecological conditions are milder or favorable (i.e., low density, high precipitation, high temperature, high beech mast), this correlation is expected to become close to zero, as even smaller offspring might be as likely to survive in such conditions. Finally, following Lack's hypothesis predicting that the size of a brood is optimised, we did not necessarily expect to find a correlation between brood size and recruitment. This is because if any deviations in brood size leads to lower recruitment, the relationship between these two variables is potentially non-linear, which cannot be captured by their correlation.

Methods

In this study, we explored whether the phenotypic correlations between three main traits linked to reproduction depended on the environmental context. To this aim, we used 58 years of

individual demographic and life-history data from the monitoring of a great tit population in Wytham Woods and applied a CRN model to brood size, offspring mass, and recruitment of offspring into the population in following years. In this model, we explored the role of ecological and environmental variables in governing the phenotypic correlations between these traits.

Study system

The great tit is a common and widespread species across Eurasia, member of the Paridae family (Gosler *et al.*, 2020), and is one of the most widely studied passerines, that has become a model species for the study of behavioral and evolutionary ecology. It dwells in woodland and wooded farmland, as well as urban areas (Gosler *et al.*, 2020), and nests in tree cavities and artificial nest boxes. Great tits are annual breeders, forming pairs and breeding in the spring, usually producing a single clutch per year consisting of 5-12 eggs, even though second clutches later in the spring sometimes occur (Perrins, 1965). The study population is from Wytham Woods (Oxfordshire, United Kingdom, 51°77'N, 1°32'W), which has been monitored since 1947 (Perrins, 1965), and the monitoring has been standardized with a stable number of nest boxes since 1961 (Perrins, 1965; Perrins & McCleery, 1989). The curated and standardized data used in this study were accessed through SPI-Birds: study name "Wytham Woods", study ID "WYT", version 1.1.0 on October 16th 2023 (Culina *et al.*, 2021). As part of this long-term individual-based monitoring, individuals were fitted with unique metal rings to allow individual identification and followed throughout their lifespan (Lack, 1964). All breeding attempts were monitored until offspring fledged, and both parents and offspring were usually caught, identified, and individually measured (e.g., mass, tarsus length).

Individual and environmental variables

Using 58 years of individual-based monitoring data from 1961 to 2018, we analyzed three different traits. These were the response variables in our model (see “Data analysis” section) and include the size of the brood (number of nestlings; measured at the brood level), the mass of offspring (g; measured at the individual offspring level), and the subsequent number of successful recruits from the brood (measured at the brood level). The mass of offspring was restricted to weight measurements of 14 days old offspring. Each brood size was recorded 14 days after hatching. We chose to analyse brood size and offspring mass instead of clutch parameters (clutch size and egg mass), which is a deviation from Lack’s principle. Brood reduction happens in our data but is infrequent and usually limited to a few offspring (Figure S1). We decided to analyse brood parameters because nestling rearing is more costly than egg production, and should better reflect the investment effort of the mothers. The fate of offspring from the brood was tracked, with recruitment defined as the number of offspring in a brood seen breeding within the population in subsequent years, a good measure of reproductive success that is often used as a proxy of parental fitness (McCleery & Clobert, 1990; Both & Visser, 2000; Wilkin *et al.*, 2006). It is important to note that both offspring mass and offspring recruitment could also be considered as offspring fitness measures rather than parental fitness (see Wolf & Wade, 2001 for a critical perspective). In addition, this measurement of offspring recruitment has often been used but is imperfect for many reasons, such as being confounded by natal dispersal. However, this bias is expected to be limited, as natal dispersal is limited in great tits, with many offspring dispersing short distances and therefore likely staying within the study population (Greenwood *et al.*, 1979; Verhulst *et al.*, 1997). In total, this yielded a final dataset encompassing 7287 broods from 5032

females across 58 years, with individual measures of mass and recruitment in the population in following years for 53753 offspring.

The explanatory variables in our model included both individual and environmental variables. The individual variables were the mass of the mother (g) of each brood, as well as the breeding age of the mother, which has an influence on reproductive success (Bouwhuis *et al.*, 2009). All breeding individuals were aged as first year breeders (1 year old) or older breeders (>1 year old), and this categorisation was based on plumage characteristics during captures when the exact age was not known from the ring number (Crates *et al.*, 2016; Simmonds *et al.*, 2020). Nestling sex is known to influence their mass, with some sexual dimorphism in body mass being present in great tit nestlings, and male nestlings weighing on average ~0.6-0.7 gram more than female nestlings around 14 days after hatching (Oddie, 2000; Tschirren *et al.*, 2003; Radersma *et al.*, 2011). However, as sex can only be determined with confidence by molecular genetic methods nestling sex was not known for most offspring, being only inferred retrospectively for the small proportion of individuals which recruited in the population in following years. Therefore, we could not include offspring sex as an explanatory variable in the model.

The environmental variables were spring temperature, spring precipitation, population density, beech mast index, and synchrony of laying dates with the caterpillar peak, which are all known to be important for great tit reproduction. These predictors were chosen because they directly influence resource availability, breeding timing, or parental effort in great tits. For instance, lower spring temperatures and precipitation may limit caterpillar abundance, a key food source, increasing the cost of brood provisioning, while population density may reflect competition for resources. Spring temperature and precipitation were obtained from the nearby

Radcliffe Meteorological Station (Burt & Burt, 2019). Spring precipitation was defined as the sum of precipitation from April 1st to May 31st, the period when offspring are in the nest (Simmonds *et al.*, 2020). Following Simmonds *et al.* (2020) analysis of climatic windows critical for great tits (Bailey & Pol, 2016; van de Pol *et al.*, 2016; Simmonds *et al.*, 2020), spring temperature was defined as the daily mean temperature from March 1st to May 9th. As usually estimated in this system, population density in a given year was calculated as the number of females hatching at least one egg (Simmonds *et al.*, 2020). Beech mast index has been routinely collected in Wytham Woods and across Europe (Lack, 1964; Perdeck *et al.*, 2000; Grøtan *et al.*, 2009), and is scored as an ordinal variable of increasing beech mast ranging from 0 to 2. Years of high beech mast are expected to be beneficial for fledgling survival in the fall and over winter by providing abundant food resources, and is therefore an important factor for offspring recruitment. Finally, synchrony is an individual's measure of phenological timing in relation to an annual, population-wide, measure of caterpillar abundance. More precisely, we calculated it as the difference between the half-fall date, which is the median date of capture of 5th instar caterpillars of the of winter moth *Opheroptera brumata* larvae, which indexes the timing of peak abundance of this species (Van Noordwijk *et al.*, 1995; Hinks *et al.*, 2015), and the female's laying date. We assessed the collinearity of all predictor variables (Figure S2), highlighting low correlations among most covariates and moderate correlation between spring temperature and population density, but collinearity of predictors is not an issue in multiple regression analyses (Morrissey & Ruxton, 2018; Vanhove, 2021).

Missing data imputation

The final datasets contained missing data for some variables. More specifically, out of a total of 7287 broods, the mass of the mother was missing for 1147 observations (15.7%), while the breeding age was not known for 210 observations (2.9%). Regarding the environmental variables across the 58 years of monitoring, 14 years had an unknown half fall date (24.1%) and 5 years had a missing beech mast index (8.6%). We imputed these missing data points using predictive mean matching with the R package *mice* (Buuren & Groothuis-Oudshoorn, 2011; Buuren *et al.*, 2023). To account for imputation uncertainty, we generated 20 alternative imputed datasets using this method, and each of these datasets was then analyzed and subsequently combined, successfully propagating imputation uncertainty in the results throughout the analysis. Additionally, we also ran the CRN model on a reduced dataset wherein the missing data were not imputed, thus keeping only complete cases, finding that parameters estimates were broadly similar in most cases (Supplementary methods; Figure S3-S4).

Data analysis

We aim to explore the ecological correlates of phenotypic correlations between maternal traits linked to reproduction in great tits, more precisely the phenotypic correlations between brood size and offspring quality (using mass as a proxy), and between offspring quality and offspring recruitment in the population in following years. To this aim, we used a multivariate Covariance Reaction Norm (CRN) model, which is a recently developed model (Martin, 2025) that we have previously tailored to routinely collected demographic data (Bliard *et al.*, 2025), allowing phenotypic covariances to vary through time or space in response to environmental variation. We encourage readers interested in the CRN methodology and its application to demographic

data to refer to Martin (2025) and Bliard et al. (2025). Here, we apply this method on the three traits of interest in this study: brood size, offspring mass, and offspring recruitment (all treated as maternal traits, analysed at the level of the brood).

We first modeled offspring mass at day 14 with a Gaussian distribution (equation 1), as a function of \mathbf{X}_1 (a $N \times P$ matrix of N measurements of P predictors, including an intercept), with predictors being spring temperature, spring precipitation, population density, beech mast index, synchrony and its quadratic effect, the mass of the mother and the age of the mother. All the continuous variables were standardized. In addition, we included a year random effect δ_1 and a nestbox random effect γ_1 . The year random effect describes the unmodelled inter-annual variability in environmental conditions, while the nestbox random effect partitions the potential variation in territory quality or non-random selection of specific nestboxes by reproductive individuals. These random effects ensure that observed phenotypic correlations are not biased by unmeasured temporal or spatial factors. We also added a year-specific mother random effect $\alpha_{1(y)}$ structured across repeated measurements by \mathbf{W} (a $N \times J$ matrix of J mothers), which as explained further below could vary in response to year-specific ecological conditions. Finally, σ is the residual variation describing the within-brood variance.

$$\mu_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \delta_1 + \gamma_1 + \mathbf{W}\boldsymbol{\alpha}_{1(y)} \quad (1)$$

$$\text{offspring mass} \sim \text{Normal}(\mu_1, \sigma_{\mu_1})$$

As our second trait, brood size, is underdispersed relative to a Poisson distribution, we modeled it using an ordinal regression (also called cumulative logistic regression; equation 2), as a function of \mathbf{X}_2 , with covariates being the same as in \mathbf{X}_1 , but not containing an intercept and differing in the number of observations. Similarly, we also included a year random effect δ_2 and a nestbox

random effect γ_2 . In the absence of repeated measurement for each mother in a given year (a single brood per female in a breeding season), as indicated by the exclusion of \mathbf{W} , we included a year-specific observation-level random effect $\mathbf{o}_{2(Y)}$, which describes the residual variation. Here, the cumulative probability of having at most i offspring is given as a function of the thresholds θ_i and the matrix of covariates \mathbf{X}_2 , as well as the random effects.

$$\text{logit}(\Pr(N_{\text{offspring}} \leq i)) = \theta_i - (\mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\delta}_2 + \gamma_2 + \mathbf{o}_{2(Y)}) \quad (2)$$

We modeled the third trait, offspring recruitment, describing the number of offspring from a given brood found breeding in the population in following years, with a zero-inflated Poisson distribution (equation 3). Recruitment data were expected to be heavily zero-inflated because many broods produced no recruits, due to factors such as brood predation and low offspring survival. The zero-inflation term ψ was included because posterior predictive checks with a simpler Poisson model without zero inflation highlighted an over-representation of zeros, causing the model to fit poorly to the data. The probability of observing a given number of offspring recruited into the population in following years was modeled as a function of \mathbf{X}_3 , which is similar to \mathbf{X}_3 with the addition of an intercept. Year $\boldsymbol{\delta}_3$ and nestbox γ_3 random effects were included as well. Since offspring recruitment was analyzed at the level of the brood (one measure of recruitment per brood), we did not have repeated measurements in a given year for mothers, and therefore included a year-specific observation-level random effect $\mathbf{o}_{3(Y)}$, characterising the residual variation.

$$\log(\mu_3) = \mathbf{X}_3\boldsymbol{\beta}_3 + \boldsymbol{\delta}_3 + \gamma_3 + \mathbf{o}_{3(Y)} \quad (3)$$

$$p(N_{\text{recruits}} | \psi, \mu_3) = \begin{cases} \psi + (1 - \psi)\text{Poisson}(0 | \mu_3), & \text{if } N_{\text{recruits}} = 0 \\ (1 - \psi)\text{Poisson}(N_{\text{recruits}} | \mu_3), & \text{if } N_{\text{recruits}} > 0 \end{cases}$$

We also tried an alternative modeling approach by using a binomial distribution to estimate the probability of offspring recruiting in the population, but unfortunately this model suffered from convergence issues that could not be resolved.

To investigate context dependence of the phenotypic correlations among the three response variables (brood size, offspring mass, offspring recruitment), the year-specific among-individual random effect $\alpha_{1(Y)}$ and observation-level random effects $\mathbf{o}_{2(Y)}$ and $\mathbf{o}_{3(Y)}$ were drawn from a multivariate normal distribution governed by year-specific covariance matrices $\mathbf{P}_{(Y)}$ (equation 4). The year-specific covariance matrices can then be decomposed in their primary elements, i.e., the year-specific phenotypic correlations between the three traits (r_{12}, r_{13}, r_{23}) and their variances ($\sigma_{\alpha_1}^2, \sigma_{o_2}^2, \sigma_{o_3}^2$), given that a covariance is just the product of the correlation between traits and the square roots of the variances. We then model the year-specific phenotypic correlations ($r_{(Y)}$), as well as the year specific variances ($\sigma^2_{(Y)}$), as a function of a subset of the environmental covariates contained in \mathbf{X}_4 (equation 4). The covariates are spring temperature, spring precipitation, population density, and beech mast index.

$$[\alpha_{1(Y)}, \mathbf{o}_{2(Y)}, \mathbf{o}_{3(Y)}] \sim \text{Multivariate Normal}(\mathbf{0}, \mathbf{P}_{(Y)}) \quad (4)$$

$$\mathbf{P}_{(Y)} = \begin{bmatrix} \sigma_{\alpha_1(Y)}^2 & r_{12(Y)} \sigma_{\alpha_1(Y)} \sigma_{o_2(Y)} & r_{13(Y)} \sigma_{\alpha_1(Y)} \sigma_{o_3(Y)} \\ r_{12(Y)} \sigma_{\alpha_1(Y)} \sigma_{o_2(Y)} & \sigma_{o_2(Y)}^2 & r_{23(Y)} \sigma_{o_2(Y)} \sigma_{o_3(Y)} \\ r_{13(Y)} \sigma_{\alpha_1(Y)} \sigma_{o_3(Y)} & r_{23(Y)} \sigma_{o_2(Y)} \sigma_{o_3(Y)} & \sigma_{o_3(Y)}^2 \end{bmatrix}$$

$$\text{atanh}(\mathbf{r}_{(Y)}) = \mathbf{X}_4 \boldsymbol{\beta}_4$$

$$\log(\sigma_{(Y)}) = \mathbf{X}_4 \boldsymbol{\beta}_5$$

Regarding the choice of priors, we used a standard flat prior for the zero-inflation coefficient ψ (equation 5), which is a default prior for this parameter in most statistical software.

$$\psi \sim \text{Beta}(1, 1) \quad (5)$$

We specified regularizing priors for all the β slope coefficients (equation 6) to reduce our risk of false positive and increase the robustness and generalizability of our findings (see Lemoine, 2019 for discussion). Narrower priors were used for β_4 to avoid putting too much weight on extreme correlations.

$$\beta_1, \beta_2, \beta_3, \beta_5 \sim \text{Normal}(0, 1) \quad (6)$$

$$\beta_4 \sim \text{Normal}(0, 0.5)$$

For the year-specific among-individual σ_α and year-specific observation-level σ_o standard deviations, as well as for the within-brood variance σ , we used exponential priors (equation 7).

$$\sigma_{\mu_1} \sim \text{Exponential}(2) \quad (7)$$

We specified the year random effects δ on the three traits studied by drawing them from univariate (i.e. non-correlated) normal distributions, also using exponential distributions for the variance terms λ (equation 8).

$$\delta_1 \sim \text{Normal}(0, \lambda_1) \quad (8)$$

$$\delta_2 \sim \text{Normal}(0, \lambda_2)$$

$$\delta_3 \sim \text{Normal}(0, \lambda_3)$$

$$\lambda_1, \lambda_2, \lambda_3 \sim \text{Exponential}(2)$$

Finally, since breeding individuals could sample nestboxes non-randomly and territories can vary in quality, the nextbox random effects γ partitioning spatial variation in the conditions experienced by individuals were specified as multivariate, thus estimating the among-nestbox covariances, as not accounting for this correlation could potentially lead to biasing the estimation of the context-dependent among-individual correlations we aimed to detect. We therefore drew the γ vectors from a multivariate normal distribution with covariance matrix Σ . The covariance matrix Σ is decomposed into its standard deviation matrix S , with exponential priors for each

standard deviation parameter ω , and its correlation matrix \mathbf{R} , which is specified with a Lewandowski-Kurowicka-Joe prior distribution (equation 9).

$$[\gamma_1, \gamma_2, \gamma_3] \sim \text{Multivariate Normal}(\boldsymbol{\Sigma}) \quad (9)$$

$$\boldsymbol{\Sigma} = \mathbf{S} \mathbf{R} \mathbf{S}$$

$$\mathbf{S} = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3 \end{bmatrix}$$

$$\omega_1, \omega_2, \omega_3 \sim \text{Exponential}(2)$$

$$\mathbf{R} \sim \text{LKJ}(2)$$

We also performed the same model as the one described above, but excluding the year random effects γ on brood size, offspring mass, and recruitment, as well as another one where we excluded both the year random effects and the environmental covariates on brood size, offspring mass, and recruitment (Figure S5-S6).

Model implementation

The multivariate CRN model (Martin, 2025; Bliard *et al.*, 2025) described above was implemented in a Bayesian framework using the statistical programming language Stan (Carpenter *et al.*, 2017), which uses an extended and cutting-edge Hamiltonian Monte Carlo algorithm (Hoffman & Gelman, 2014). We ran the model through R version 4.3 (R Core Team, 2023), using the R package CmdStanR version 0.8.1 (Gabry & Češnovar, 2020). As detailed in the previous methods section, weakly informative regularizing priors were used for all parameters. We performed 20 alternative models, one for each of the imputed dataset, to ensure that imputation uncertainty was propagated through the analysis. Each model ran on 3 chains with a warm-up period of 1000 iterations, and sampled for 1000 iterations per chain, keeping all the iterations (Link & Eaton, 2012), thus totalling 3000 saved posterior samples per alternative model. We ensured that

convergence was reached for each model by obtaining R-hat values below 1.01 for all parameters (Gelman & Rubin, 1992), and by visually inspecting trace plots. We then merged the posterior distributions of all the 20 alternative models, thus obtaining a single posterior distribution made of 60000 posterior samples (3000 x 20) for the combined model accounting for imputation uncertainty. Throughout the results, we report the posterior median effect sizes, alongside credible intervals.

Results

Our results reveal that the correlation between brood size and offspring mass is negative across all environmental contexts, with a slight increase under favourable conditions. In contrast, the phenotypic correlation between offspring mass and recruitment is strongly context-dependent, becoming indistinguishable from zero under favourable conditions such as low density but persisting in harsher years. Contrary to our hypothesis, we found no correlation between brood size and recruitment.

We expected the correlation between brood size and offspring mass to be highly constrained, whereby mothers producing larger broods would also produce smaller offspring. The results of our model potentially confirm this expectation, as the correlation between brood size residual variation and among-mother variation in offspring mass was found to be negative across all environmental contexts (Figure 1). The negative correlation between brood size and offspring mass was slightly relaxed in years of high precipitation, and, to a lesser extent, in years of low population density, although these effects were small and are associated with high uncertainties (Figure 1).

The correlation between the mass of offspring produced and their recruitment in subsequent years tended to be positive overall (Figure 2), such that mothers producing larger offspring tend to have higher offspring recruitment. This follows expectations that larger offspring are more likely to survive, thus being more likely to be present as breeders in following years. However, we found rather strong effects of the environmental context on this correlation. Following our expectations, under favorable conditions such as low population density, high spring temperature, or high spring precipitation, this correlation tended to be small, such that the mass of offspring was not clearly associated with their recruitment (Figure 2). While we expected the same for beech mast index, we found the opposite result, with the correlation becoming indistinguishable from zero in years of low beech mast (Figure 2).

Contrary to our hypothesis, we found no correlation between brood size and recruitment, and this correlation was invariant across environmental contexts (Figure 3).

The among-nestbox correlations between brood size and offspring mass (median = 0.033 [10-90% intervals = -0.519; 0.563]) and between brood size and recruitment (0.032 [-0.525; 0.575]) were found to be close to zero. However, we found an overall positive correlation among nest boxes between offspring mass and recruitment (0.387 [-0.249; 0.776]), which could reflect some degree of spatial variation in nestbox or overall territory quality.

Estimated effects of all the covariates on the three phenotypic correlations highlight that the uncertainty around the median estimated effects is fairly large despite the high sample size used (Figure 4), as we are ultimately limited by the dataset length in the number of breeding seasons monitored. In addition, it is important to note that some results are sensitive to model structure, with slightly different estimates found when year random effects were excluded

(Figure S5, S6). Environmental covariates are also found to influence trait variances in various ways (Figure S9, S10, S11, S12). Overall, covariate effects on the primary traits (Figure S7, S8) align with previous studies: population density negatively affected reproductive traits, while beech mast index and maternal breeding age had positive effects across all traits. In addition, posterior predictive checks highlight a good fit of the model to the data (Figure S13).

Discussion

We used individual-based data from one of the longest individual-based monitoring of a great tit population, together with a newly developed hierarchical “covariance reaction norm” model (Martin, 2025) tailored for demographic data lacking repeated measurements within years for some traits (Bliard *et al.*, 2025). Despite some uncertainty in the results, we found evidence that the phenotypic correlations between brood size and offspring mass, and especially between offspring mass and offspring recruitment, are temporally variable and dependent on the environmental conditions experienced during the breeding season. Overall, harsher conditions lead females who lay larger broods to have smaller offspring, and the propensity for these smaller offspring to recruit is lower. Interestingly, smaller offspring born during favorable breeding seasons are equally likely to recruit than their larger counterparts. Yet, regardless of the environment, the fitness outcome for mothers does not seem to be influenced by the size of their brood, as it is not ultimately associated with offspring recruitment, though offspring recruitment is not always a reliable fitness proxy for mothers (Wolf & Wade, 2001). We discuss these findings and potential limitations of our modeling approach that could blur some of the biological associations studied here.

The correlation between offspring quantity and offspring quality was found to be strongly negative and highly constrained in this population of great tit. Regardless of the environmental conditions experienced during the breeding seasons, the correlation between brood size residual variation and the among-female offspring mass variation was close to -1, thus being potentially indicative of a strong quantity-quality trade-off. We only found a weak influence of the environmental context (Figure 4), though these effects were statistically uncertain and sensitive to data imputation (Figure S3, S4) and model structure (Figure S5, S6). Convincing evidence of context-dependence between offspring number and size has been found in studies on human fertility, whereby such trade-off is absent in favorable socio-economic classes while being found under poorer socio-economic conditions (Gillespie *et al.*, 2008; Meij *et al.*, 2009; Lawson & Mulder, 2016). Similarly, laboratory studies on invertebrates found phenotypic and genetic correlations between offspring number and size to depend on the degree of food availability (Messina & Slade, 1999; Czesak & Fox, 2003; Messina & Fry, 2003). However, we found that the correlation is negative across environments, potentially indicating that this trade-off is always expressed in great tits, and a potential explanation could be that little variance in quality or resource acquisition remains among females once accounting for primary predictors such as mother's mass and age. For instance, the study of Ebert (1993) on *Daphnia* has shown that the offspring number and size trade-off is initially found to be influenced by food availability, but the genetic correlations all became negative once accounting for mother's condition. Similarly, negative phenotypic correlations between offspring quantity and quality were found after adjusting for maternal size in a meta-analysis across animal species (Lim *et al.*, 2014). Therefore,

if our primary covariates accounted well for maternal heterogeneity in this great tit population, it could result in such a strongly negative correlation.

We found that the correlation between offspring mass and recruitment is overall positive (Figure 2), such that larger offspring are better quality offspring, thus being more likely to recruit in the population in following years. This is in line with numerous other past results in great tits, wherein offspring mass or size are usually found to be associated with future outcomes, from survival to recruitment (Both *et al.*, 1999; Monrós *et al.*, 2002; Garant *et al.*, 2004; Wilkin *et al.*, 2006; Bouwhuis *et al.*, 2015; Rodríguez *et al.*, 2016), while being independent of their laying date (Wilkin *et al.*, 2006). However, our study also explores the dependence of this association on environmental conditions experienced during the breeding season. Interestingly, we found that this association was stronger under harsh conditions, whereby producing small offspring might be particularly detrimental as they would not fare well due to poor climatic conditions (e.g., dry and cold springs; Figure 2) or stronger competition from a higher population density (Both *et al.*, 1999). However, during favorable breeding seasons, the correlation between offspring mass and recruitment becomes indistinguishable from zero, with mild conditions and low competition allowing even frail offspring to survive and recruit in following seasons, thus highlighting that phenotypic selection for offspring body mass is likely variable and its temporal dynamics are potentially mediated by environmental conditions (Braby, 1994; Grant & Grant, 2002; Siepielski *et al.*, 2009). This echoes findings from Bouwhuis *et al.* (2015), where broods with heavier offspring experienced a better recruitment probability, with this relationship being stronger during warmer springs. Such fluctuating selection is expected to promote the evolution of adaptive plasticity in reproductive behavior (de Jong, 1995). Despite most results following our

expectations, the estimated effect for the beech mast index goes opposite to our predictions, with theoretically better years (i.e., higher beech mast index) associated with a stronger, more positive correlation. Although speculative, one possible explanation could be that females do an anticipatory adjustment of brood size based on future resource availability, whereby they would adjust brood size based on some external cues that correlate with beech seed production in the autumn following the breeding season. For instance, these anticipatory effects in relation to masting events have been found in red squirrels *Tamiasciurus hudsonicus* and *Sciurus vulgaris* (Boutin *et al.*, 2006). In great tits, this might lead to increased brood size at the population level during years of high beech mast (Figure S7), which in turn is likely to increase nestling and fledgling abundance in the population. This increase in the abundance of great tits and potentially other bird species might lead to a decrease in individual perceived predation risk (e.g., through dilution effects), thus making it less detrimental for nestlings to put on weight (Gentle & Gosler, 2001; Macleod *et al.*, 2005). Though the estimated effect is rather modest in size and somewhat uncertain (Figure 4), warranting caution with biological interpretation, this highlights the need for further exploration of the indirect effects of environmental variables on reproductive trade-offs.

Overall, we found a lack of correlation between brood size and subsequent successful recruitment, together with no evidence of a dependence on the environmental context. Several non-mutually exclusive reasons, both biological and statistical, could explain this invariant correlation, indistinguishable from zero across all environments. First, the lack of association could just reflect that after accounting for primary predictors on both traits, there is just no biological association between residual brood size and recruitment and no environmental

effects. This may simply be due to the fact that residual variances for size and particularly recruitment were very small (Fig. S10-11), limiting our ability to estimate their correlation independently of mean-scaling effects, despite our large sample size. However, this correlation did not change even when excluding primary predictors on traits (Figure S6). This lack of association would mean that regardless of environmental conditions, producing a small brood of big nestlings or a large brood of small nestlings is likely leading to the same fitness outcome for females. Second, the lack of correlation between brood size and offspring recruitment could actually reflect Lack's initial statement, with brood size being optimized (Williams, 1966). Indeed, under this hypothesis, we might expect a non-linear association between brood size and fitness, such that recruitment initially increases with brood size up to an optimal value, and then decreases as brood size effects on offspring quality become apparent. Therefore, any deviation is likely leading to reduced recruitment (Gustafsson & Sutherland, 1988; Pettifor *et al.*, 1988). While the covariance reaction norm model is ideal for estimating nonlinear changes in trait associations, this non-linear relationship would not be directly captured by the covariance terms estimated in our model, which could explain the absence of a clear correlation in our results. These non-linear associations could be more directly analysed using recently proposed methods for nonlinear selection analysis on latent variables (Dingemanse *et al.*, 2021; Martin *et al.*, 2021), but this would require a highly complex and much more difficult to interpret hierarchical model in the context of this study. Third, brood size and recruitment are both non-repeated measures within a given year (i.e., context). Thus, by not having several measures per individual in a given year, within- and among-individual covariances cannot be properly disentangled. Such observation-level correlation will therefore reflect either the among-individual correlation if trait

505 repeatabilities are high, or the within-individual correlation if trait repeatabilities are low (Bliard
506 *et al.*, 2025). The repeatability of clutch size in great tits is usually medium to high (e.g., 0.51 in
507 Perrins and Jones (1974)). However, if females' offspring recruitment is environmentally labile, it
508 would have a low repeatability, and our estimate may mostly reflect the within-individual
509 correlation between both traits (Searle, 1961; Dingemanse & Dochtermann, 2013; Bliard *et al.*,
510 2025).

511 Other limitations could explain the lack of correlation between brood size and
512 recruitment. Recruitment of offspring into the population in following years is an imperfect
513 measure of fitness. As previously noted, because offspring effects on recruitment may be
514 independent of maternal traits, it might thus be a poor proxy of females' fitness per se (Wolf &
515 Wade, 2001). It is necessarily an underestimate due to imperfect detection, being confounded
516 with long distance (i.e., outside the study area) natal dispersal (Gimenez *et al.*, 2008). Dispersal
517 events could themselves be linked to the environmental context experienced by the nestlings
518 (McCaslin *et al.*, 2020). For instance, it had been argued that social dominance, which could
519 hypothetically be related to the size of the brood an offspring was reared in, could in turn
520 influence the natal dispersal distance of offspring (Nilsson & Smith, 1985; Smith & Nilsson, 1987;
521 Smith *et al.*, 1989). Such a limitation could potentially obscure any association between brood
522 size and recruitment (Gimenez *et al.*, 2008). Finally, it could also be possible that the lack of
523 association and lack of environmental effects result from interaction effects between
524 environmental variables that were not accounted for, even though we did not have specific
525 biologically motivated interaction to include.

Altogether, we found indication that, although the brood size – offspring mass trade-off was highly constrained, the phenotypic correlation between offspring mass and recruitment was strongly dependent on changes in population density and harshness of the environment during the breeding season, with females producing larger offspring experiencing higher recruitment only during harsh breeding seasons. This study demonstrates that in this great tit population, phenotypic covariances respond to continuous environmental change and temporal variation in population density, a phenomenon that has mostly been neglected up to now, which highlights that knowledge about life-history theory and trade-offs can be improved through the incorporation of context dependence. Future studies could extend this approach to other populations or species (Culina *et al.*, 2021), to further understand how environmental variation shapes trade-offs and life-history evolution (Chantepie *et al.*, 2024). Overall, such a framework allows us to revisit old ecological questions related to patterns of selection in fluctuating density-dependent environments through the lens of new multivariate statistical methods, and therefore expand from an often-univariate view on the topic to studying multivariate patterns of trait (co)expression (Wright *et al.*, 2019; Martin, 2025).

Acknowledgements

This work was supported by a Swiss National Science Foundation Grant (31003A_182286 to A.O). We acknowledge the many hundreds of people who have collected data as part of the long-term study in Wytham. Recent work on the Wytham long-term study has been funded by grants from NERC (NE/X000184/1, NE/K006274/1, NE/S010335/1), BBSRC (BB/L006081/1), the ERC (AdG250164), and UKRI (EP/X024520/1). We are grateful to SPI-Birds Network and Database and

its members for all their efforts through the years to assemble this network, thus facilitating access to the data used in this manuscript. Finally, we thank two anonymous reviewers for their constructive comments and suggestions.

Authors contributions

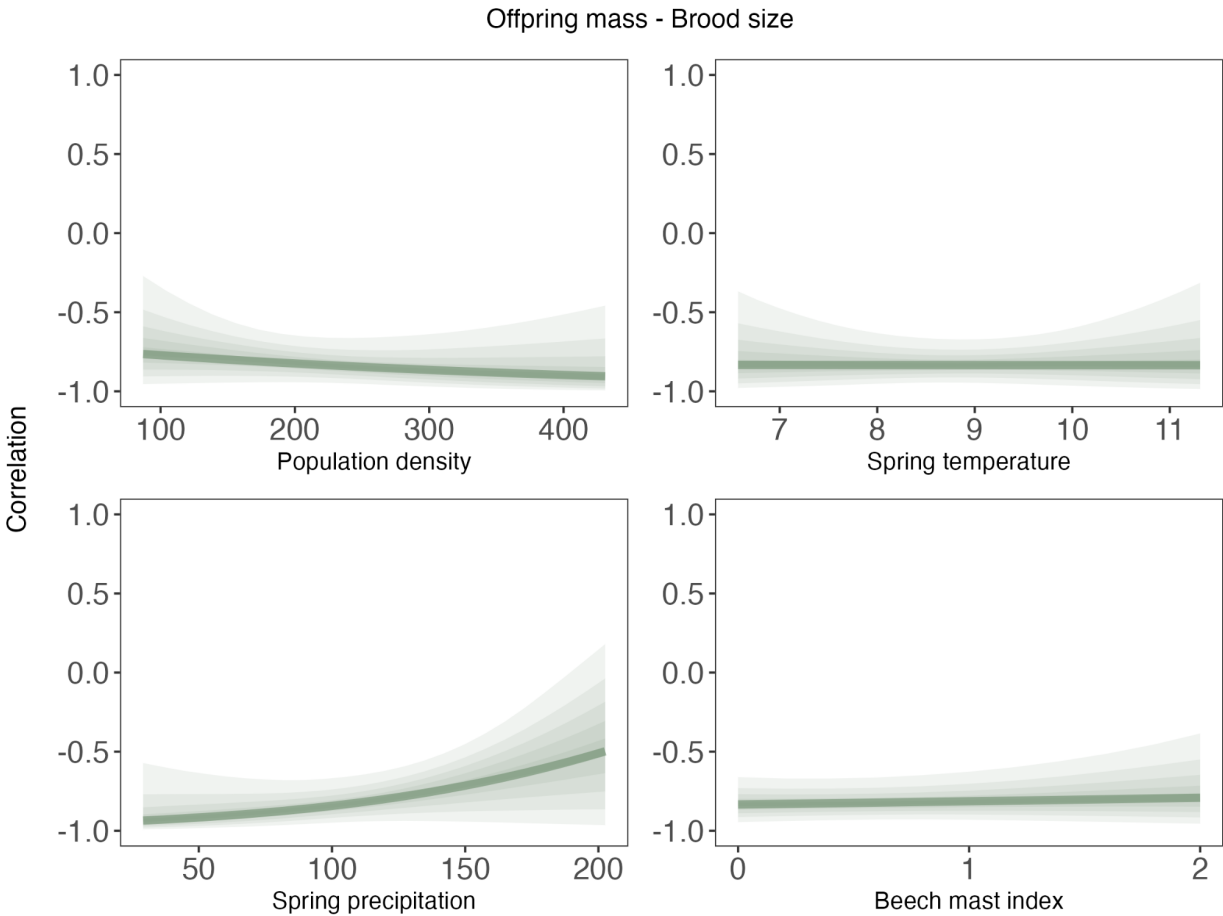
LB conceived the study and analyzed the data, using an initial modeling framework developed by JSM. EC and BS collected and curated the data. LB wrote the first draft. All authors contributed to the editing of the manuscript.

Data and code availability

The formatted data, as well as the R and Stan code necessary to reproduce the results are available on GitHub https://github.com/lbiard/tradeoffs_parus_major and will be archived on Zenodo. The raw datasets analyzed in the current study are available in the SPI-Birds Database (study name: Wytham Woods, study ID: WYT, version: 1.0.0, data custodian: Ben Sheldon).

570

Figures



571

572 Figure 1: Estimated phenotypic correlations between offspring mass and brood size as a function
573 of population size (top left panel), spring temperature (top right panel), spring precipitation
574 (bottom left panel), and beech mast index (bottom right panel). Posterior median effect sizes are
575 represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded
576 bands.

577

578

579

580

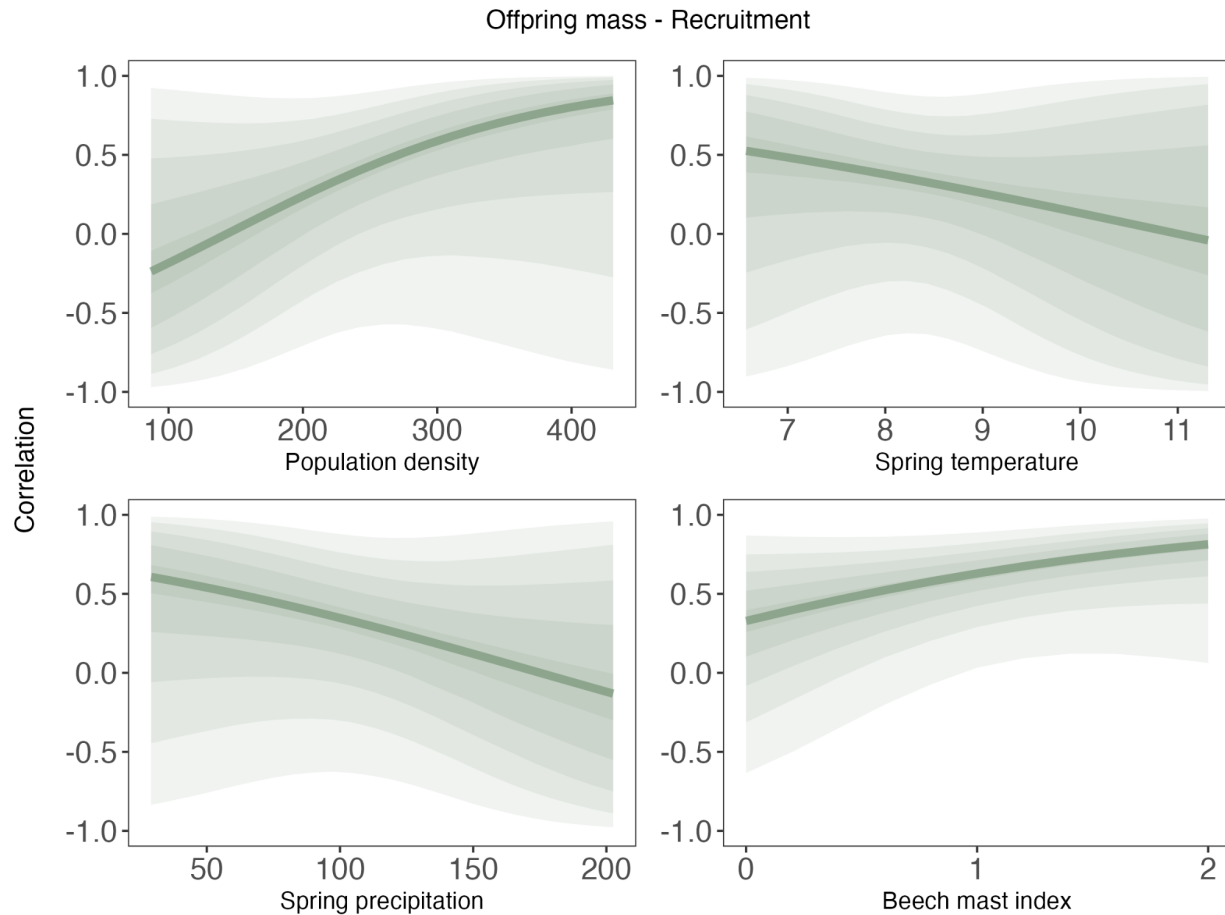


Figure 2: Estimated phenotypic correlations between offspring mass and recruitment as a function of population size (top left panel), spring temperature (top right panel), spring precipitation (bottom left panel), and beech mast index (bottom right panel). Posterior median effect sizes are represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded bands.

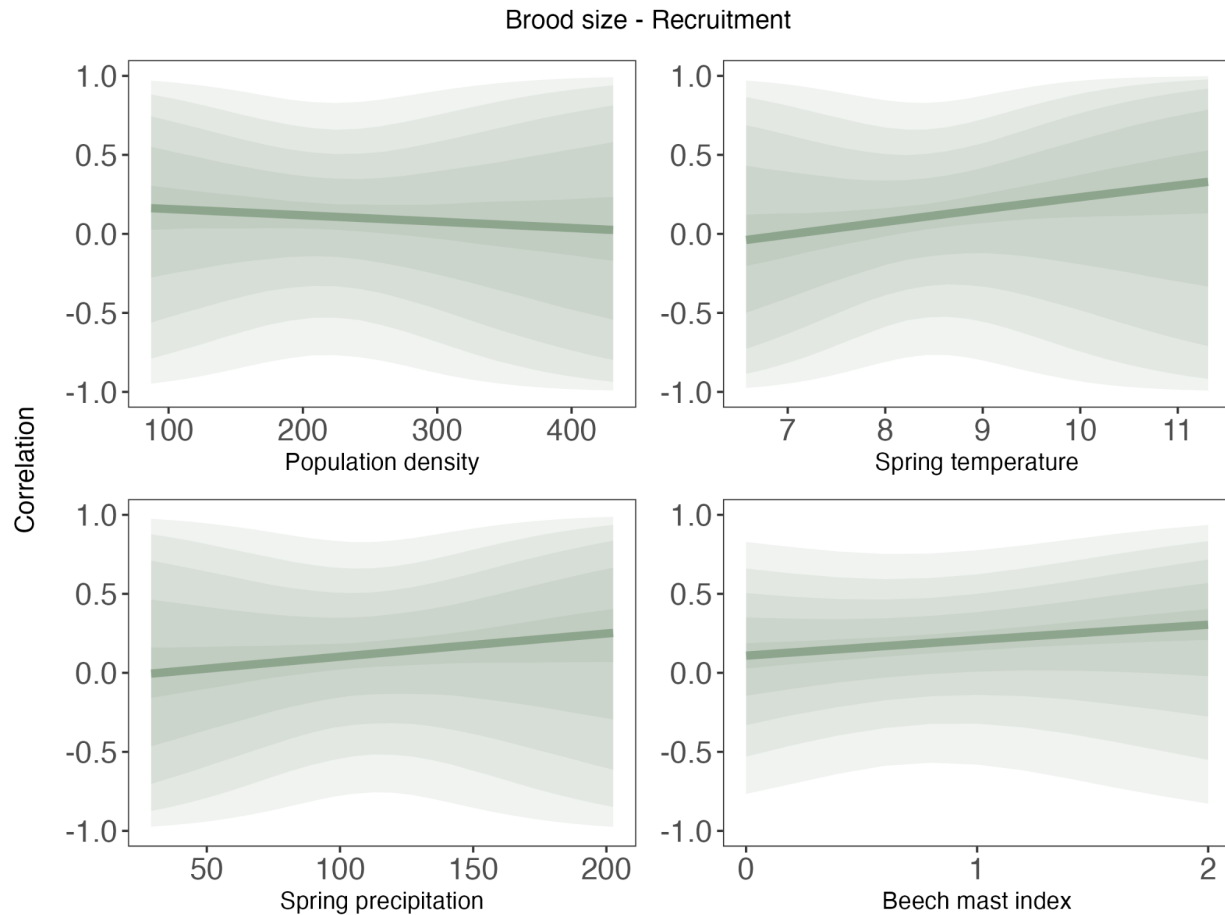


Figure 3: Estimated phenotypic correlations between brood size and recruitment as a function of population size (top left panel), spring temperature (top right panel), spring precipitation (bottom left panel), and beech mast index (bottom right panel). Posterior median effect sizes are represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded bands.

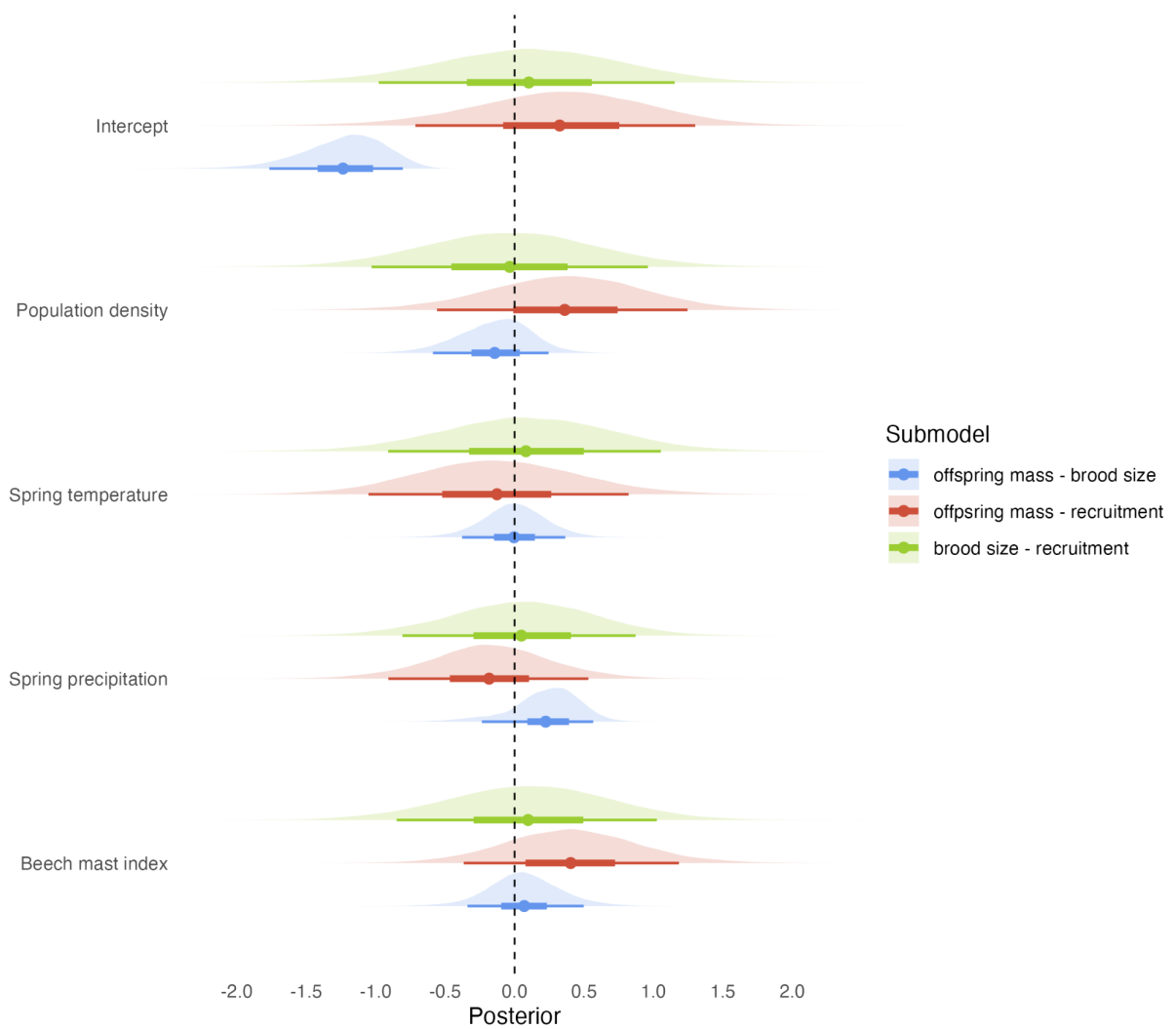


Figure 4: Estimated effects of standardized predictors on the phenotypic correlations between offspring mass and brood size (blue), offspring mass and recruitment (red), and brood size and recruitment (green).

Supplementary materials

Supplementary methods:

As mentioned in the main text, the final dataset contained missing data for some variables in some observations. We imputed these missing data points using predictive mean matching with the R package *mice*. For each missing entry, the method compares this observation with all other complete cases, and based on the other variables selects a group of candidate observations (5 observations, the default in *mice* R package) that most closely matches the given observation with missing data. From this pool of 5 observations, one is then drawn at random and used to impute the missing data. Given that it draws real values from the data, it is less likely to impute implausible data, and the method usually works well (Kleinke 2017).

To assess the influence of the data imputation, we performed the same CRN model keeping only complete cases, thus excluding any observation that contained missing data for at least one variable. This yielded a smaller dataset made of 35570 offspring mass measurements from 4710 breeding attempts. However, this also led to a quite strong reduction in the number of breeding seasons with complete case observations, with this reduced dataset spanning only 27 breeding seasons (vs. 58 breeding seasons in the full dataset). Overall, we found broad concordance between the results of the CRN model with full imputed dataset and reduced non-imputed dataset. However, we want to point out that some specific parameter estimates appeared to change between the two analyses (Figure S3; effect of population density and spring precipitation on the brood size - offspring mass correlation). But given that these effects did not involve any of the imputed variables, it is more likely that these changes stem from the loss of more than half of the breeding seasons in the reduced dataset (58 vs 27 years).

Supplementary figures:

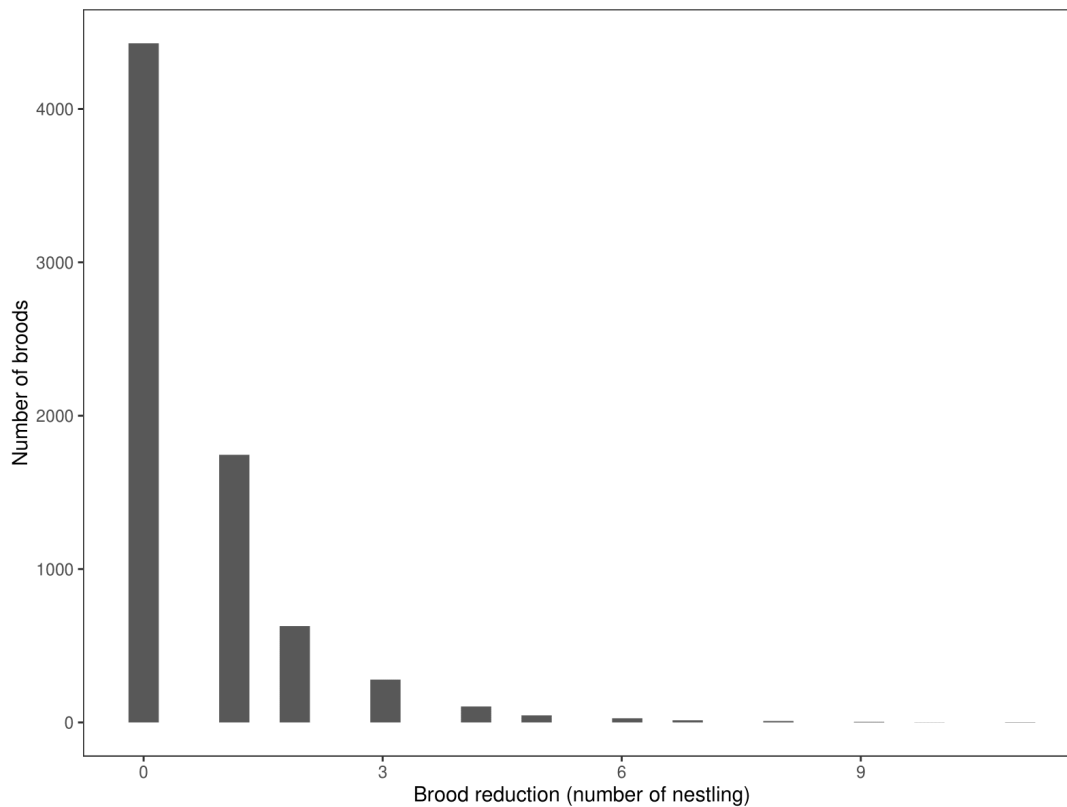


Figure S1: histogram of brood reduction across all broods analysed. This highlights the number of broods analysed that went through a brood reduction, with values representing the difference between observed clutch size and observed brood size at day 14. Brood reduction appears to be infrequent or limited to a few offspring.

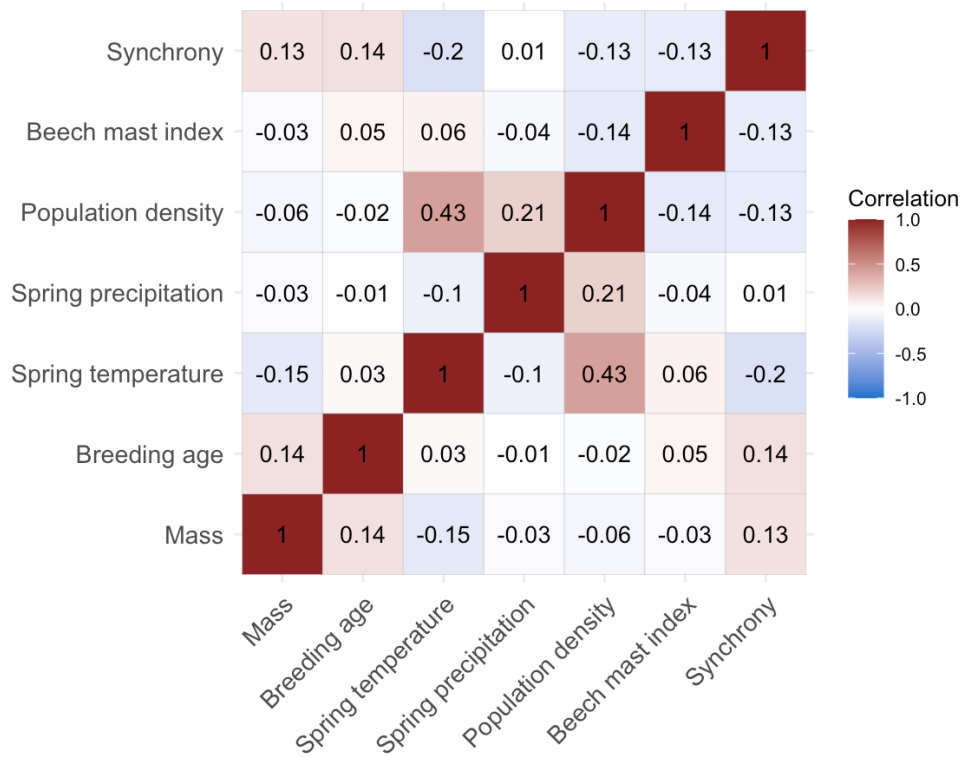


Figure S2: Correlations among the predictor variables used in the model, highlighting low to moderate collinearity among all variables.

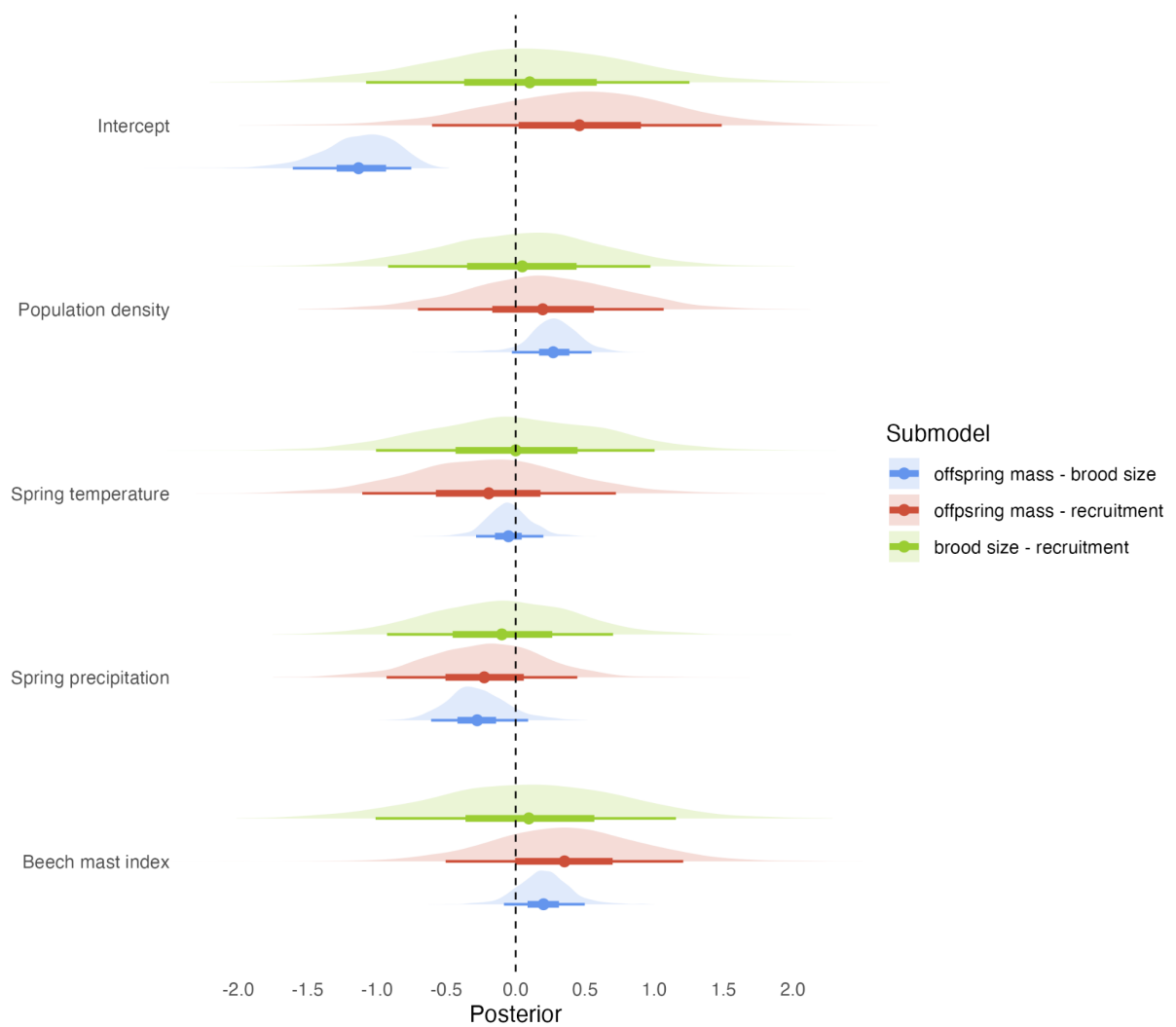


Figure S3: Estimated effects of standardized predictors on the phenotypic correlations between offspring mass and brood size (blue), offspring mass and recruitment (red), and brood size and recruitment (green), using the reduced, non-imputed dataset.

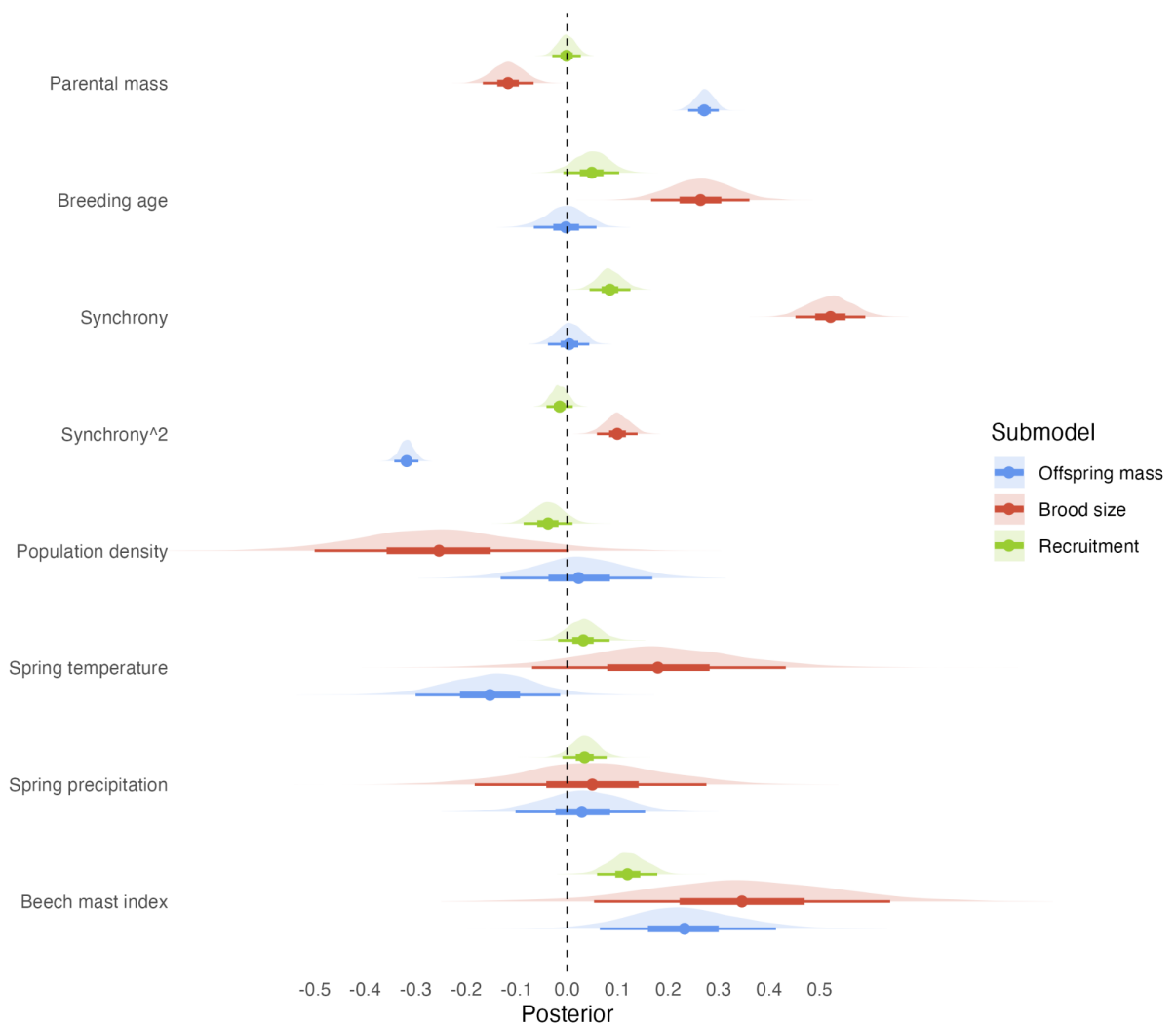


Figure S4: Estimated effects of standardized predictors on primary traits: offspring mass (blue), brood size (red), and recruitment (green), using the reduced, non-imputed dataset.

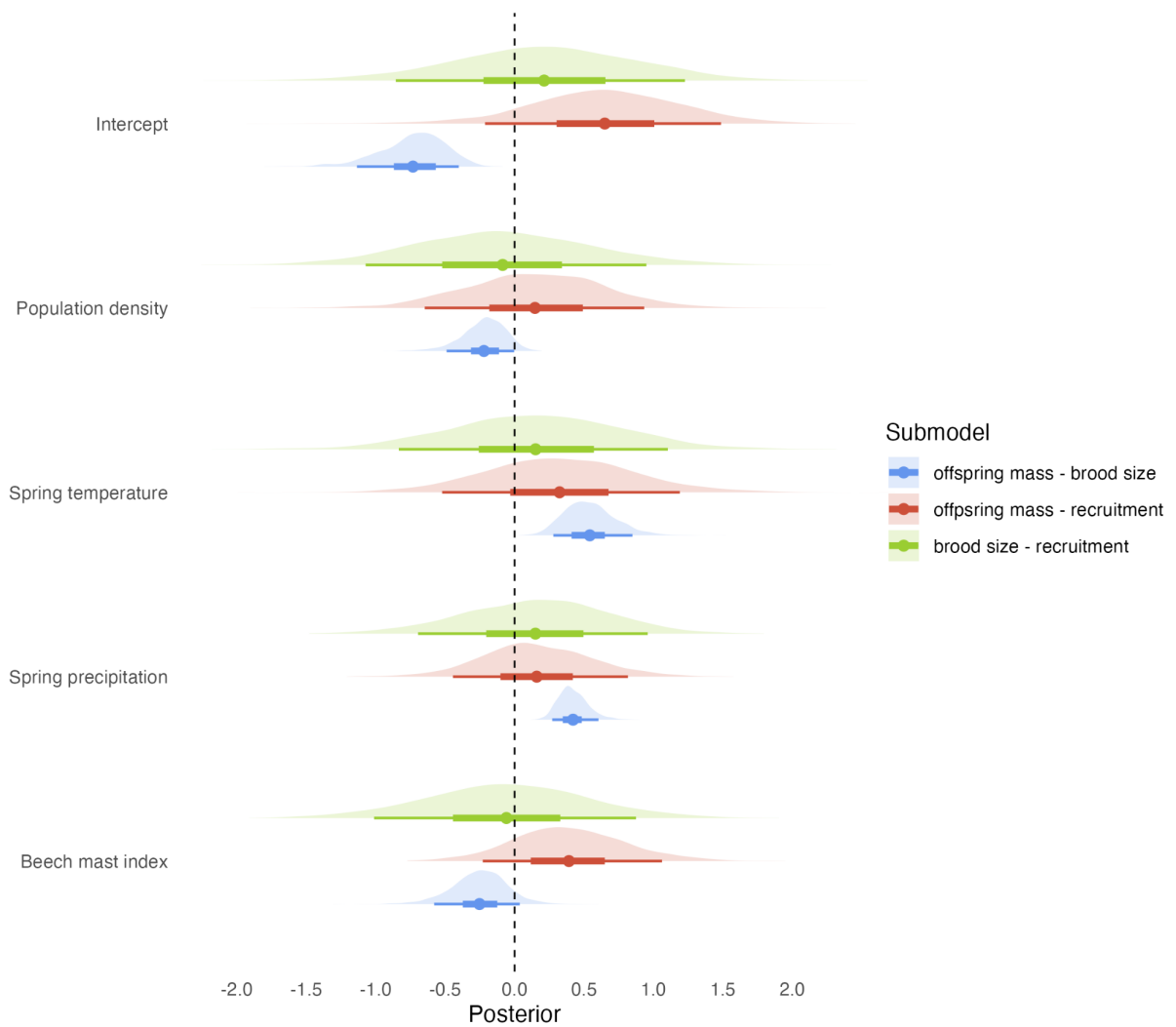


Figure S5: Estimated effects of standardized predictors on the phenotypic correlations between offspring mass and brood size (blue), offspring mass and recruitment (red), and brood size and recruitment (green), when using a model that did not include the year random effects on primary traits.

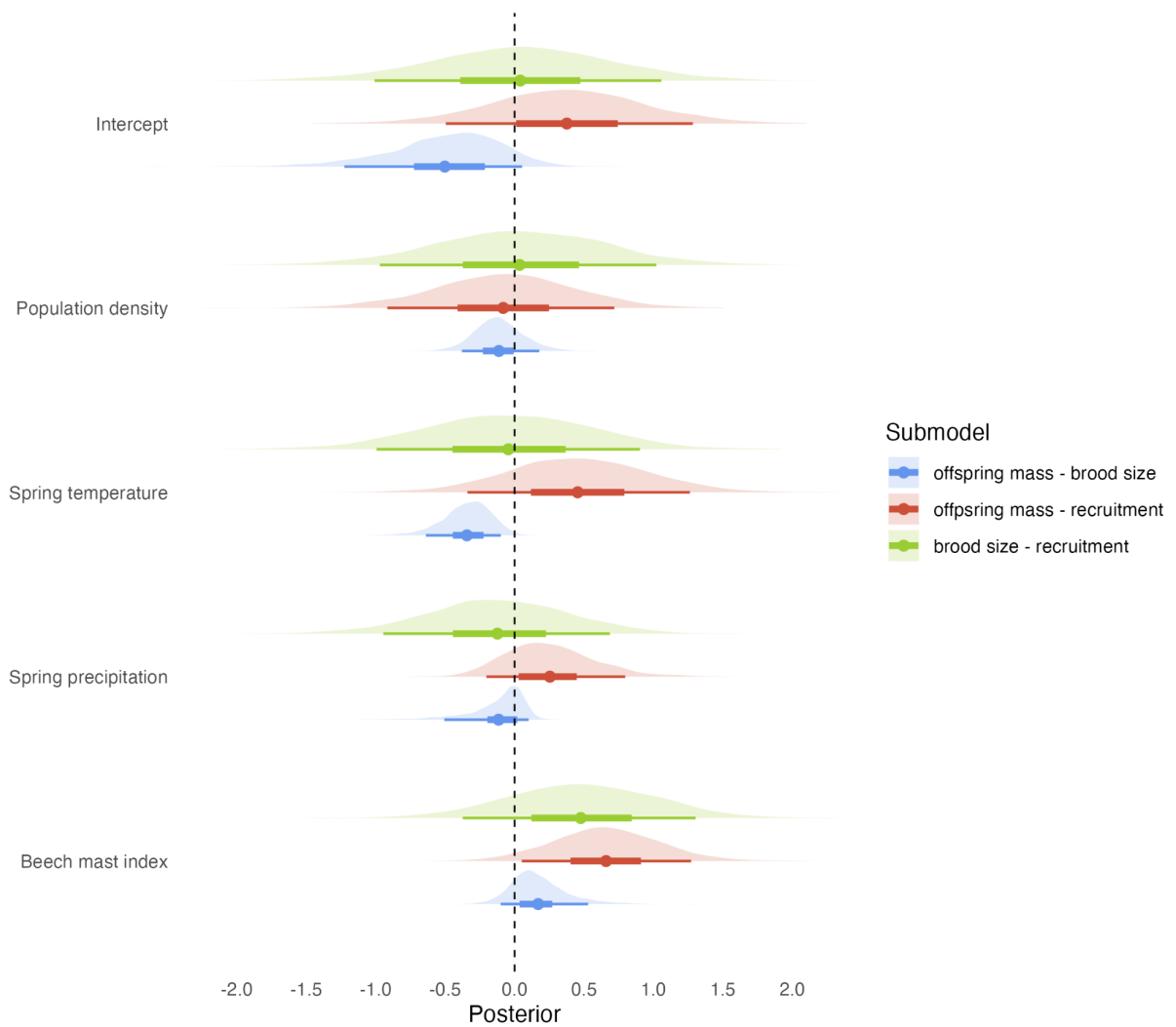


Figure S6: Estimated effects of standardized predictors on the phenotypic correlations between offspring mass and brood size (blue), offspring mass and recruitment (red), and brood size and recruitment (green), when using a model that did not include the year random effects on primary traits, nor the environmental covariates on the primary traits.

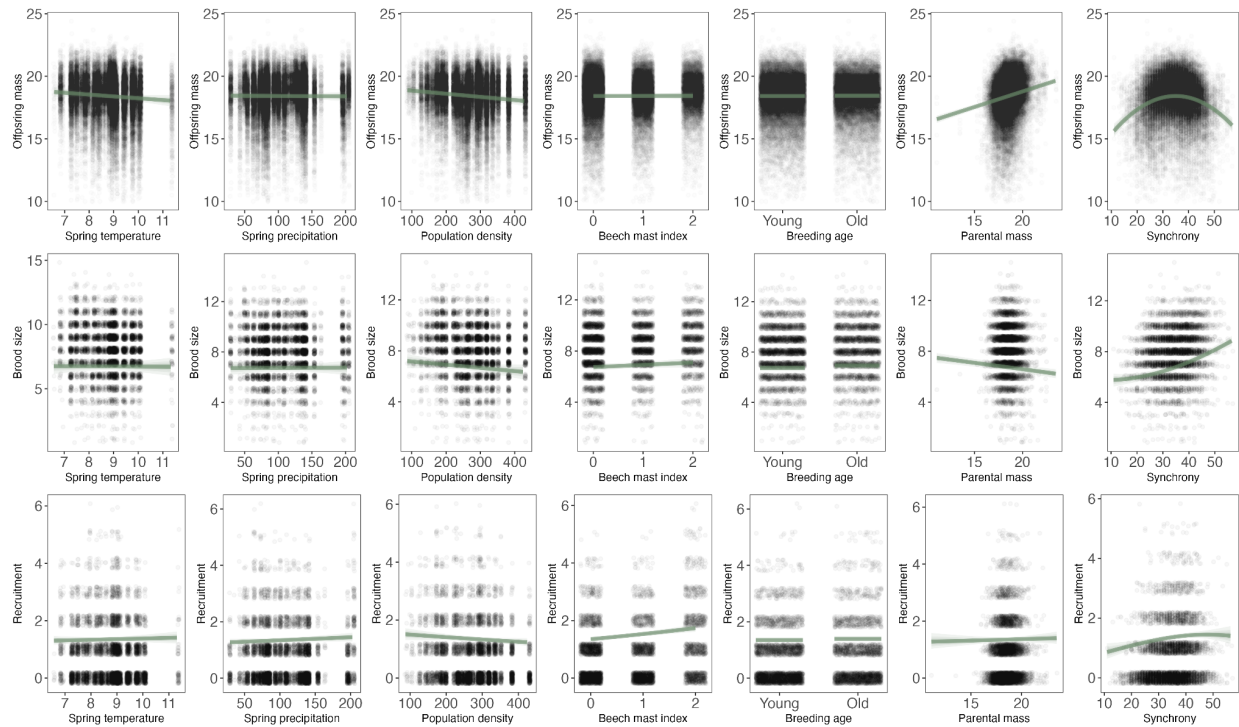


Figure S7: Associations estimated by the model between spring temperature, spring precipitation, population density, beech mast index, breeding age, parental mass, and laying date synchrony (panels from left to right) with offspring mass (top row), brood size (middle row), and recruitment (bottom row). Posterior median effect sizes are represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded bands. Each circle represents a data point, and a small amount of jitter and transparency were added for display only to avoid overlapping points.

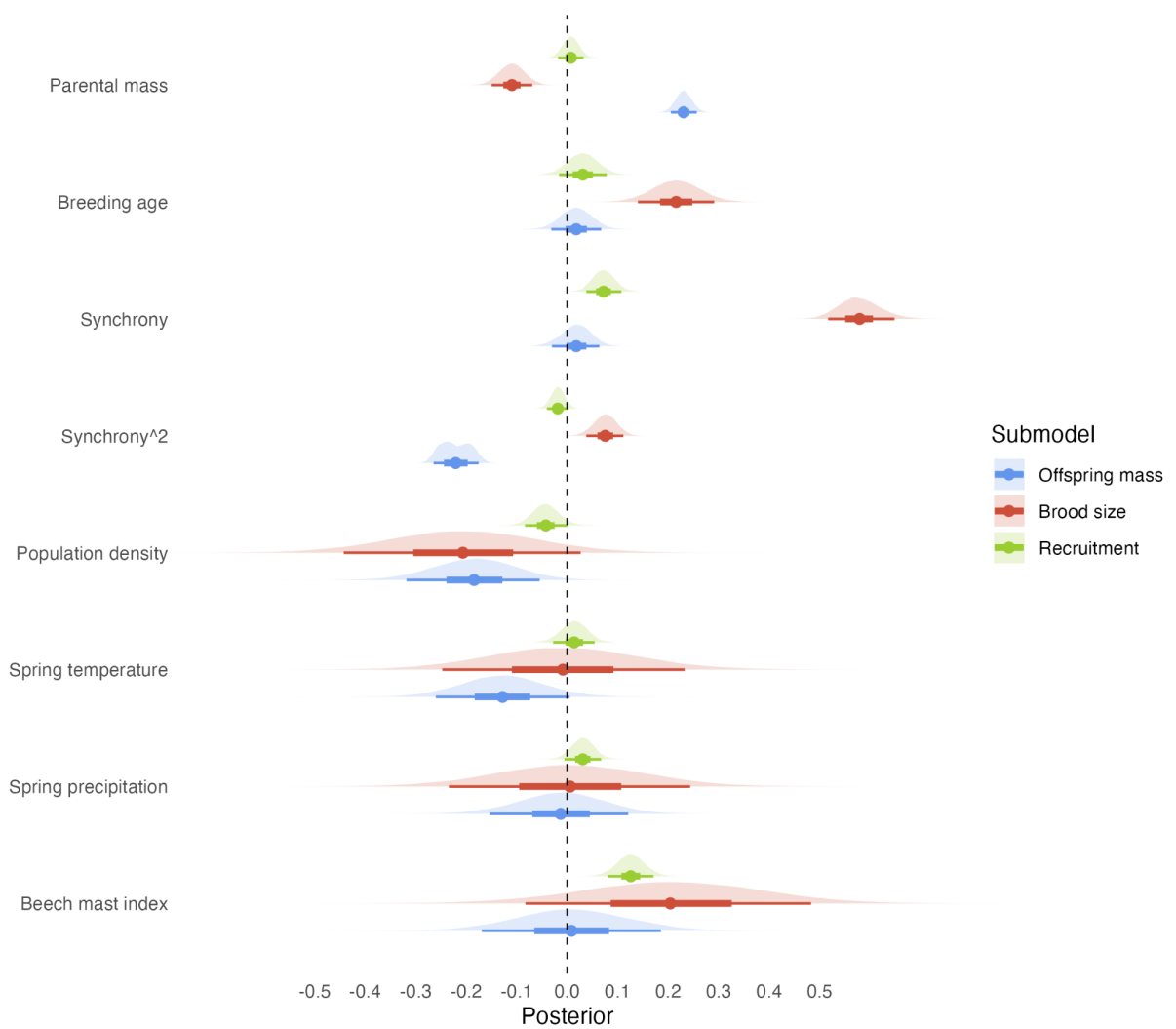


Figure S8: Estimated effects of standardized predictors on primary traits: offspring mass (blue), brood size (red), and recruitment (green).

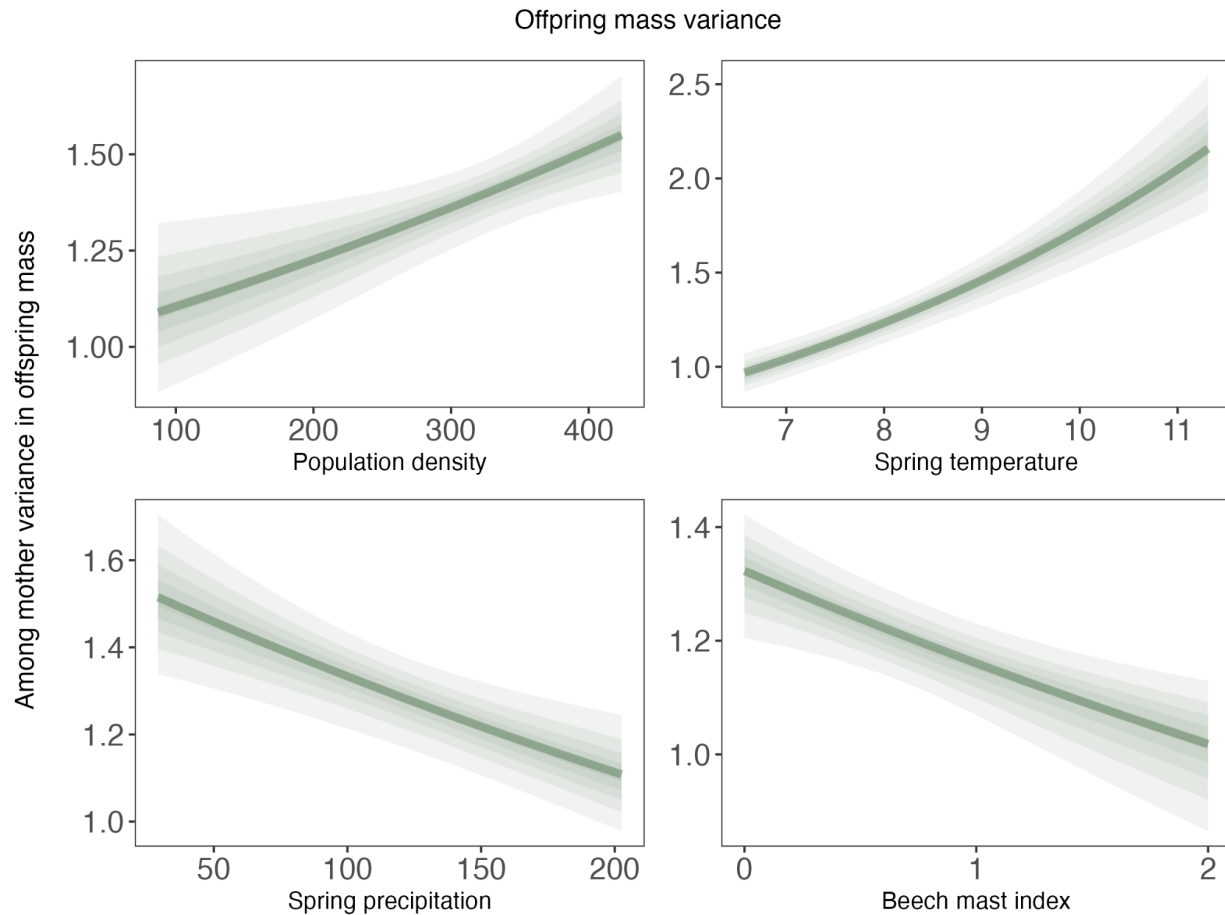


Figure S9: Estimated among-mother variance in offspring mass as a function of population size (top left panel), spring temperature (top right panel), spring precipitation (bottom left panel), and beech mast index (bottom right panel). Posterior median effect sizes are represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded bands.

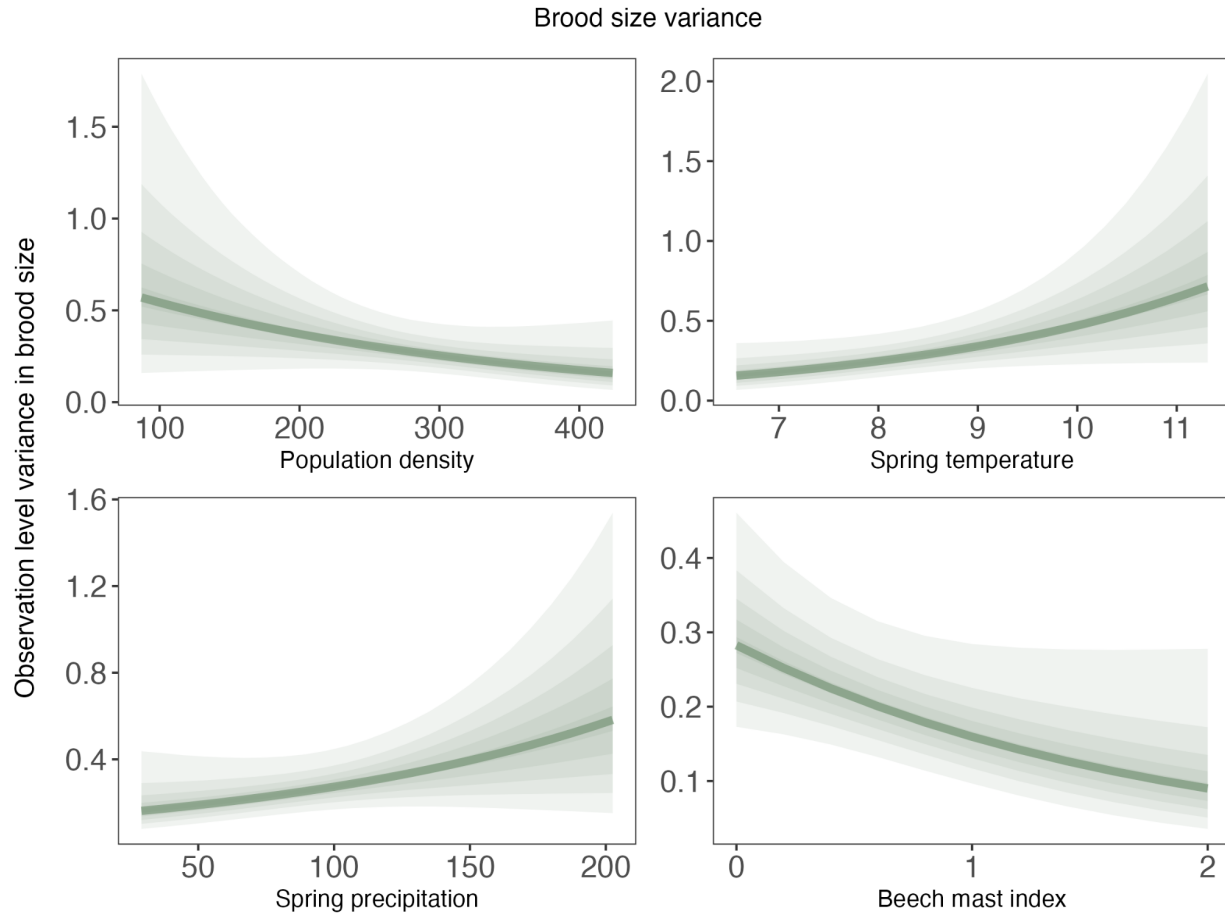


Figure S10: Estimated observation-level variance in brood size as a function of population size (top left panel), spring temperature (top right panel), spring precipitation (bottom left panel), and beech mast index (bottom right panel). Posterior median effect sizes are represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded bands.

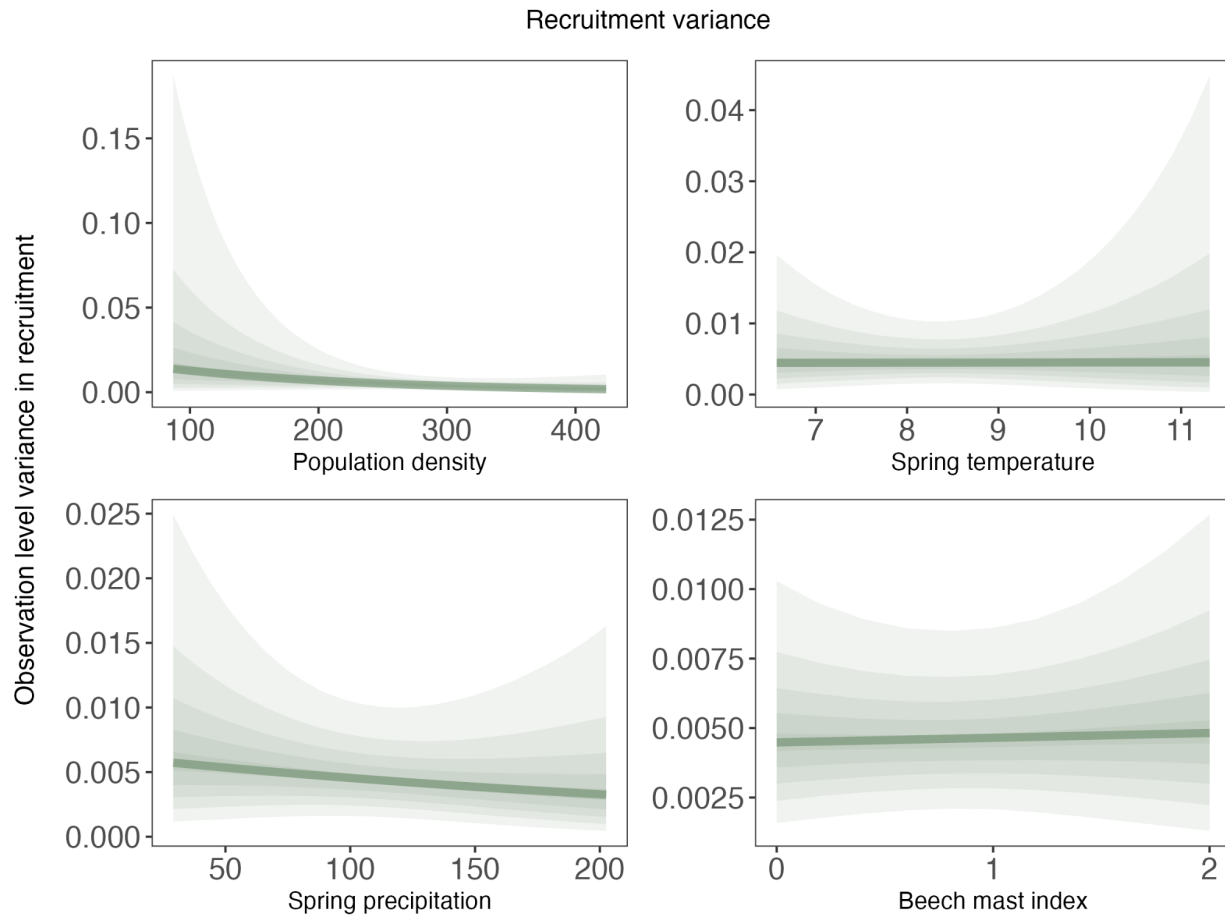


Figure S11: Estimated observation-level variance in offspring recruitment as a function of population size (top left panel), spring temperature (top right panel), spring precipitation (bottom left panel), and beech mast index (bottom right panel). Posterior median effect sizes are represented by the darker lines, and 10% to 90% credible intervals are represented by the shaded bands.

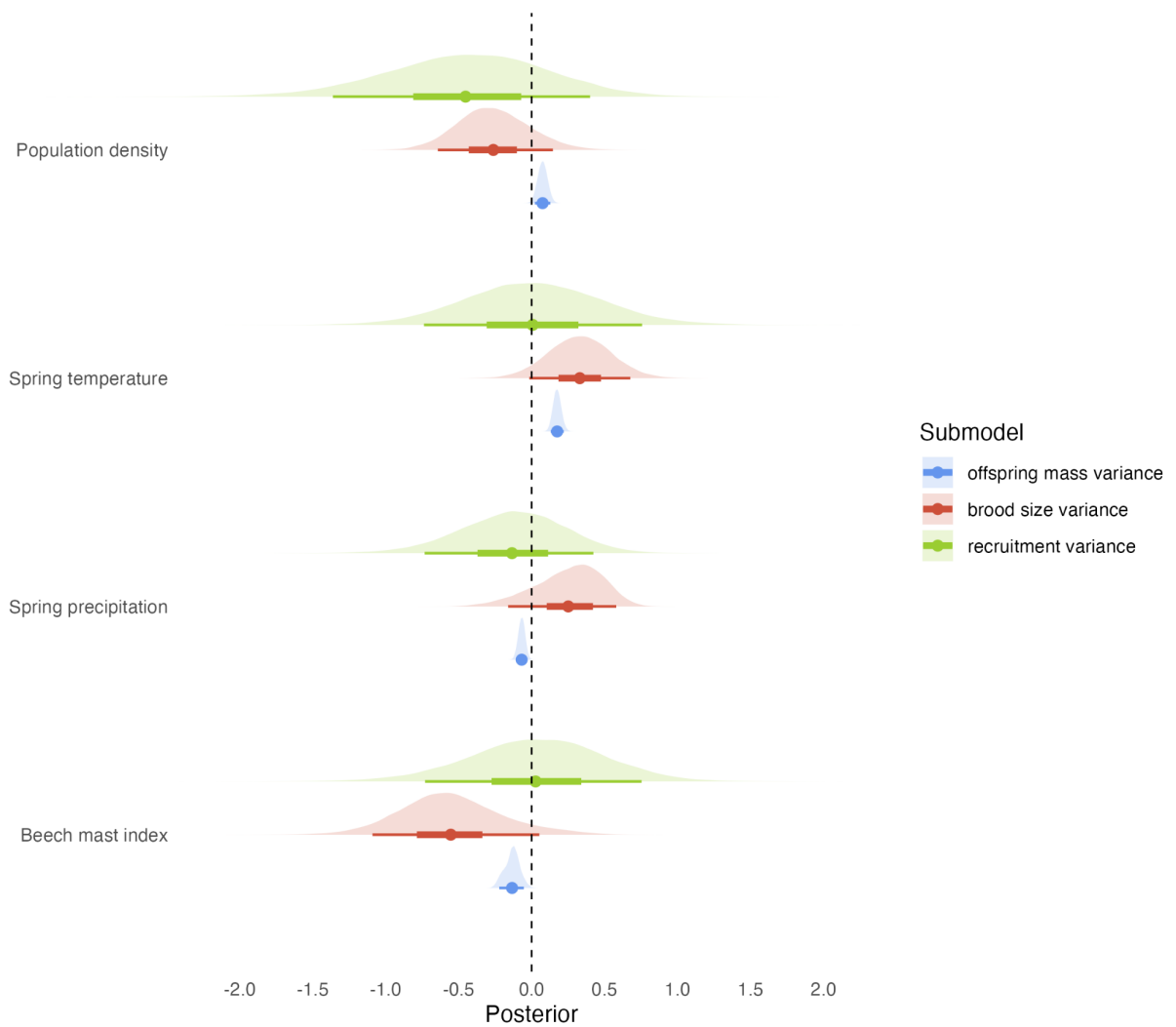


Figure S12: Estimated effects of standardized predictors on the among-mother variance in offspring mass (blue), observation-level variance in brood size (red), and observation-level variance in offspring recruitment (green).

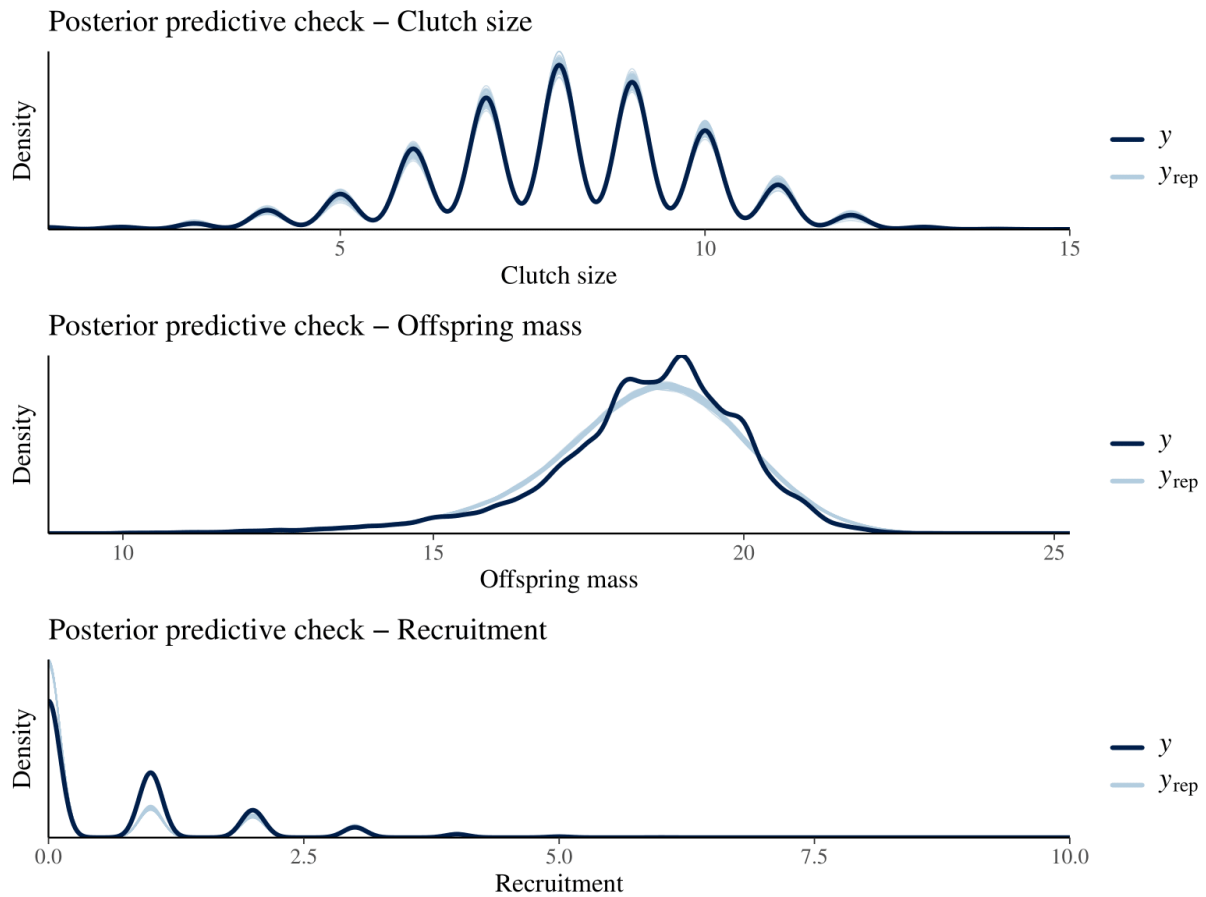


Figure S13: Posterior predictive checks showing the concordance between the distribution of the data (y) and the distribution of data generated under the statistical model (y_{rep}), for brood size (top panel), offspring mass (middle panel), and offspring recruitment (bottom panel). This highlights a decent fit of the models, but a small overestimation of zeros in recruitment.

759

References

760

Allen, R.M., Buckley, Y.M. & Marshall, D.J. 2008. Offspring Size Plasticity in Response to

761

Intraspecific Competition: An Adaptive Maternal Effect across Life-History Stages. *The*

762

American Naturalist **171**: 225–237. The University of Chicago Press.

763

Bailey, L.D. & Pol, M. van de. 2016. climwin: An R Toolbox for Climate Window Analysis. *PLOS*

764

ONE **11**: e0167980. Public Library of Science.

765

Beltran, R.S. & Tarwater, C.E. 2024. Overcoming the pitfalls of categorizing continuous variables

766

in ecology, evolution and behaviour. *Proceedings of the Royal Society B: Biological*

767

Sciences **291**: 20241640. Royal Society.

768

Björklund, M. 2004. Constancy of the G matrix in ecological time. *Evolution* **58**: 1157–1164.

769

Bliard, L., Martin, J.S., Paniw, M., Blumstein, D.T., Martin, J.G.A., Pemberton, J.M., *et al.* 2025.

770

Detecting context dependence in the expression of life history trade-offs. *Journal of*

771

Animal Ecology **94**: 379–393.

772

Both, C. & Visser, M.E. 2000. Breeding territory size affects fitness: an experimental study on

773

competition at the individual level. *Journal of Animal Ecology* **69**: 1021–1030.

774

Both, C., Visser, M.E. & Verboven, N. 1999. Density-dependent recruitment rates in great tits:

775

the importance of being heavier. *Proceedings of the Royal Society of London. Series B:*

776

Biological Sciences **266**: 465–469. Royal Society.

777

Boutin, S., Wauters, L.A., McAdam, A.G., Humphries, M.M., Tosi, G. & Dhondt, A.A. 2006.

778 Anticipatory Reproduction and Population Growth in Seed Predators. *Science* **314**:
779 1928–1930. American Association for the Advancement of Science.

780 Bouwhuis, S., Sheldon, B. c., Verhulst, S. & Charmantier, A. 2009. Great tits growing old:
781 selective disappearance and the partitioning of senescence to stages within the
782 breeding cycle. *Proceedings of the Royal Society B: Biological Sciences* **276**: 2769–2777.
783 Royal Society.

784 Bouwhuis, S., Vedder, O., Garroway, C.J. & Sheldon, B.C. 2015. Ecological causes of multilevel
785 covariance between size and first-year survival in a wild bird population. *Journal of*
786 *Animal Ecology* **84**: 208–218.

787 Boyce, M.S. & Perrins, C.M. 1987. Optimizing Great Tit Clutch Size in a Fluctuating Environment.
788 *Ecology* **68**: 142–153.

789 Braby, M.F. 1994. The Significance of Egg Size Variation in Butterflies in Relation to Hostplant
790 Quality. *Oikos* **71**: 119–129. [Nordic Society Oikos, Wiley].

791 Burt, S. & Burt, T. 2019. *Oxford Weather and Climate since 1767*. Oxford University Press.

792 Buuren, S. van & Groothuis-Oudshoorn, K. 2011. mice: Multivariate Imputation by Chained
793 Equations in R. *Journal of Statistical Software* **45**: 1–67.

794 Buuren, S. van, Groothuis-Oudshoorn, K., Vink, G., Schouten, R., Robitzsch, A., Rockenschaub,
795 P., *et al.* 2023. mice: Multivariate Imputation by Chained Equations.

796 Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., *et al.* 2017.

797 Stan: A Probabilistic Programming Language. *Journal of Statistical Software* **76**: 1–32.

798 Chantepie, S., Charmantier, A., Delahaie, B., Adriaensen, F., Matthysen, E., Visser, M.E., *et al.*
799 2024. Divergence in evolutionary potential of life history traits among wild populations
800 is predicted by differences in climatic conditions. *Evolution Letters* **8**: 29–42.

801 Cohen, A.A., Coste, C.F.D., Li, X.-Y., Bourg, S. & Pavard, S. 2020. Are trade-offs really the key
802 drivers of ageing and life span? *Functional Ecology* **34**: 153–166.

803 Crates, R.A., Firth, J.A., Farine, D.R., Garroway, C.J., Kidd, L.R., Aplin, L.M., *et al.* 2016. Individual
804 variation in winter supplementary food consumption and its consequences for
805 reproduction in wild birds. *Journal of Avian Biology* **47**: 678–689.

806 Culina, A., Adriaensen, F., Bailey, L.D., Burgess, M.D., Charmantier, A., Cole, E.F., *et al.* 2021.
807 Connecting the data landscape of long-term ecological studies: The SPI-Birds data hub.
808 *Journal of Animal Ecology* **90**: 2147–2160.

809 Czesak, M.E. & Fox, C.W. 2003. Evolutionary Ecology of Egg Size and Number in a Seed Beetle:
810 Genetic Trade-Off Differs Between Environments. *Evolution* **57**: 1121–1132.

811 de Jong, G. 1995. Phenotypic Plasticity as a Product of Selection in a Variable Environment. *The*
812 *American Naturalist* **145**: 493–512. [The University of Chicago Press, The American
813 Society of Naturalists].

814 de Jong, G. 1989. Phenotypically Plastic Characters in Isolated Populations. In: *Evolutionary*
815 *Biology of Transient Unstable Populations* (A. Fontdevila, ed), pp. 3–18. Springer, Berlin,

816 Heidelberg.

817 Dingemanse, N.J., Araya-Ajoy, Y.G. & Westneat, D.F. 2021. Most published selection gradients
818 are underestimated: Why this is and how to fix it. *Evolution* **75**: 806–818.

819 Dingemanse, N.J. & Dochtermann, N.A. 2013. Quantifying individual variation in behaviour:
820 mixed-effect modelling approaches. *Journal of Animal Ecology* **82**: 39–54.

821 Ebert, D. 1993. The trade-off between offspring size and number in *Daphnia magna*: the
822 influence of genetic, environmental and maternal effects. *Arch. Hydrobiol.* **90**: 453–473.

823 Einum, S. & Fleming, I.A. 2000. Highly fecund mothers sacrifice offspring survival to maximize
824 fitness. *Nature* **405**: 565–567. Nature Publishing Group.

825 Fischer, B., Taborsky, B. & Kokko, H. 2011. How to balance the offspring quality–quantity
826 tradeoff when environmental cues are unreliable. *Oikos* **120**: 258–270.

827 Fox, C.W. & Czesak, M.E. 2000. Evolutionary Ecology of Progeny Size in Arthropods. *Annual*
828 *Review of Entomology* **45**: 341–369. Annual Reviews.

829 Gabry, J. & Češnovar, R. 2020. cmdstanr: R Interface to “CmdStan.”

830 Garant, D., Kruuk, L.E.B., McCleery, R.H. & Sheldon, B.C. 2004. Evolution in a Changing
831 Environment: A Case Study with Great Tit Fledging Mass. *The American Naturalist* **164**:
832 E115–E129. The University of Chicago Press.

833 Gelman, A. & Rubin, D.B. 1992. Inference from Iterative Simulation Using Multiple Sequences.

834 *Statistical Science* **7**: 457–472. Institute of Mathematical Statistics.

835 Gentle, L.K. & Gosler, A.G. 2001. Fat reserves and perceived predation risk in the great tit, *Parus*
836 *major*. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**:
837 487–491. Royal Society.

838 Gillespie, D.O.S., Russell, A.F. & Lummaa, V. 2008. When fecundity does not equal fitness:
839 evidence of an offspring quantity versus quality trade-off in pre-industrial humans.
840 *Proceedings of the Royal Society B: Biological Sciences* **275**: 713–722. Royal Society.

841 Gimenez, O., Viallefont, A., Charmantier, A., Pradel, R., Cam, E., Brown, C.R., *et al.* 2008. The
842 Risk of Flawed Inference in Evolutionary Studies When Detectability Is Less than One.
843 *The American Naturalist* **172**: 441–448. The University of Chicago Press.

844 Gosler, A., Clement, P. & Christie, D. 2020. Great Tit (*Parus major*), version 1.0. *Birds of the*
845 *World*, doi: 10.2173/bow.gretit1.01species_shared.bow.project_name. Cornell Lab of
846 Ornithology, Ithaca, NY, USA.

847 Grant, P.R. & Grant, B.R. 2002. Unpredictable Evolution in a 30-Year Study of Darwin’s Finches.
848 *Science* **296**: 707–711. American Association for the Advancement of Science.

849 Greenwood, P.J., Harvey, P.H. & Perrins, C.M. 1979. The role of dispersal in the great tit (*Parus*
850 *major*): the causes, consequences and heritability of natal dispersal. *Journal of Animal*
851 *Ecology* **48**: 123–142. [Wiley, British Ecological Society].

852 Grøtan, V., Sæther, B.-E., Engen, S., Van Balen, J.H., Perdeck, A.C. & Visser, M.E. 2009. Spatial

853 and temporal variation in the relative contribution of density dependence, climate
854 variation and migration to fluctuations in the size of great tit populations. *Journal of*
855 *Animal Ecology* **78**: 447–459.

856 Gustafsson, L. & Sutherland, W.J. 1988. The costs of reproduction in the collared flycatcher
857 *Ficedula albicollis*. *Nature* **335**: 813–815.

858 Hinks, A.E., Cole, E.F., Daniels, K.J., Wilkin, T.A., Nakagawa, S. & Sheldon, B.C. 2015. Scale-
859 Dependent Phenological Synchrony between Songbirds and Their Caterpillar Food
860 Source. *The American Naturalist* **186**: 84–97. The University of Chicago Press.

861 Hoffman, M.D. & Gelman, A. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in
862 Hamiltonian Monte Carlo. *Journal of Machine Learning Research* **15**: 1593–1623.

863 Högstedt, G. 1980. Evolution of Clutch Size in Birds: Adaptive Variation in Relation to Territory
864 Quality. *Science* **210**: 1148–1150. American Association for the Advancement of Science.

865 Houslay, T.M., Vierbuchen, M., Grimmer, A.J., Young, A.J. & Wilson, A.J. 2018. Testing the
866 stability of behavioural coping style across stress contexts in the Trinidadian guppy.
867 *Functional Ecology* **32**: 424–438.

868 Lack, D. 1964. A Long-Term Study of the Great Tit (*Parus major*). *Journal of Animal Ecology* **33**:
869 159–173. [Wiley, British Ecological Society].

870 Lack, D. 1954. *The natural regulation of animal numbers*. Oxford : Clarendon Press.

871 Lawson, D.W. & Mulder, M.B. 2016. The offspring quantity–quality trade-off and human fertility

872 variation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**:
873 20150145. Royal Society.

874 Leips, J., Richardson, J.M.L., Rodd, F.H. & Travis, J. 2009. Adaptive Maternal Adjustments of
875 Offspring Size in Response to Conspecific Density in Two Populations of the Least
876 Killifish, *Heterandria Formosa*. *Evolution* **63**: 1341–1347.

877 Lemoine, N.P. 2019. Moving beyond noninformative priors: why and how to choose weakly
878 informative priors in Bayesian analyses. *Oikos* **128**: 912–928.

879 Lim, J.N., Senior, A.M. & Nakagawa, S. 2014. Heterogeneity in individual quality and
880 reproductive trade-offs within species. *Evolution* **68**: 2306–2318.

881 Link, W.A. & Eaton, M.J. 2012. On thinning of chains in MCMC. *Methods in Ecology and*
882 *Evolution* **3**: 112–115.

883 MacCallum, R.C., Zhang, S., Preacher, K.J. & Rucker, D.D. 2002. On the practice of
884 dichotomization of quantitative variables. *Psychological Methods* **7**: 19–40. American
885 Psychological Association, US.

886 Macleod, R., Gosler, A.G. & Cresswell, W. 2005. Diurnal mass gain strategies and perceived
887 predation risk in the great tit *Parus major*. *Journal of Animal Ecology* **74**: 956–964.

888 Marshall, D.J., Bolton, T.F. & Keough, M.J. 2003. Offspring Size Affects the Post-Metamorphic
889 Performance of a Colonial Marine Invertebrate. *Ecology* **84**: 3131–3137. Ecological
890 Society of America.

891 Marshall, D.J., Bonduriansky, R. & Bussière, L.F. 2008. Offspring Size Variation Within Broods as
892 a Bet-Hedging Strategy in Unpredictable Environments. *Ecology* **89**: 2506–2517.

893 Martin, J.S. 2025. Covariance reaction norms: A flexible method for estimating complex
894 environmental effects on trait (co)variances. *Methods in Ecology and Evolution* **00**: 1-20.

895 Martin, J.S., Araya-Ajoy, Y.G., Dingemanse, N.J., Wilson, A.J. & Westneat, D. 2021. Estimating
896 (non)linear selection on reaction norms: A general framework for labile traits.
897 EcoEvoRxiv.

898 Martin, J.S., Westneat, D., Nakagawa, S., Schroeder, J. & Burke, T. 2025. Measuring
899 selection on reaction norms: Lack’s principle and plasticity in clutch size. *Evolution*
900 qpaf168.

901 McCaslin, H.M., Caughlin, T.T. & Heath, J.A. 2020. Long-distance natal dispersal is relatively
902 frequent and correlated with environmental factors in a widespread raptor. *Journal of*
903 *Animal Ecology* **89**: 2077–2088.

904 McCleery, R.H. & Clobert, J. 1990. Differences in Recruitment of Young by Immigrant and
905 Resident Great Tits in Wytham Wood. In: *Population Biology of Passerine Birds*, pp. 423–
906 440. Springer, Berlin, Heidelberg.

907 Meij, J.J., Van Bodegom, D., Ziem, J.B., Amankwa, J., Polderman, A.M., Kirkwood, T.B.L., *et al.*
908 2009. Quality–quantity trade-off of human offspring under adverse environmental
909 conditions. *Journal of Evolutionary Biology* **22**: 1014–1023.

910 Messina, F.J. & Fry, J.D. 2003. Environment-dependent reversal of a life history trade-off in the
 911 seed beetle *Callosobruchus maculatus*. *Journal of Evolutionary Biology* **16**: 501–509.

912 Messina, F.J. & Slade, A.F. 1999. Expression of a life-history trade-off in a seed beetle depends
 913 on environmental context. *Physiological Entomology* **24**: 358–363.

914 Mitchell, D.J. & Houslay, T.M. 2021. Context-dependent trait covariances: how plasticity shapes
 915 behavioral syndromes. *Behavioral Ecology* **32**: 25–29.

916 Monrós, J.S., Belda, E.J. & Barba, E. 2002. Post-fledging survival of individual great tits: the
 917 effect of hatching date and fledging mass. *Oikos* **99**: 481–488.

918 Morrissey, M.B. & Ruxton, G.D. 2018. Multiple Regression Is Not Multiple Regressions: The
 919 Meaning of Multiple Regression and the Non-Problem of Collinearity. *Philosophy,*
 920 *Theory, and Practice in Biology* **10**.

921 Nilsson, J.-Å. & Smith, H.G. 1985. Early Fledgling Mortality and the Timing of Juvenile Dispersal
 922 in the Marsh Tit *Parus palustris*. *Ornis Scandinavica (Scandinavian Journal of*
 923 *Ornithology)* **16**: 293–298. [Nordic Society Oikos, Wiley].

924 Nur, N. 1984. The Consequences of Brood Size for Breeding Blue Tits II. Nestling Weight,
 925 Offspring Survival and Optimal Brood Size. *Journal of Animal Ecology* **53**: 497–517.
 926 [Wiley, British Ecological Society].

927 Oddie, K.R. 2000. Size matters: competition between male and female great tit offspring.
 928 *Journal of Animal Ecology* **69**: 903–912.

929 Parker, G.A. & Begon, M. 1986. Optimal Egg Size and Clutch Size: Effects of Environment and
930 Maternal Phenotype. *The American Naturalist* **128**: 573–592. The University of Chicago
931 Press.

932 Perdeck, A., Visser, M. & Balen, J. 2000. Great Tit *Parus major* survival, and the beech-crop
933 cycle. *Ardea* **88**: 99–108.

934 Perrins, C.M. 1965. Population Fluctuations and Clutch-Size in the Great Tit, *Parus major* L.
935 *Journal of Animal Ecology* **34**: 601–647. [Wiley, British Ecological Society].

936 Perrins, C.M. & Jones, P.J. 1974. The Inheritance of Clutch Size in the Great Tit (*Parus Major* L.).
937 *The Condor* **76**: 225–229.

938 Perrins, C.M. & McCleery, R.H. 1989. Laying Dates and Clutch Size in the Great Tit. *The Wilson*
939 *Bulletin* **101**: 236–253. Wilson Ornithological Society.

940 Perrins, C.M. & Moss, D. 1975. Reproductive Rates in the Great Tit. *Journal of Animal Ecology*
941 **44**: 695–706. [Wiley, British Ecological Society].

942 Pettifor, R.A., Perrins, C.M. & McCleery, R.H. 1988. Individual optimization of clutch size in great
943 tits. *Nature* **336**: 160–162. Nature Publishing Group.

944 Pettifor, R.A., Perrins, C.M. & McCleery, R.H. 2001. The individual optimization of fitness:
945 variation in reproductive output, including clutch size, mean nestling mass and offspring
946 recruitment, in manipulated broods of great tits *Parus major*. *Journal of Animal Ecology*
947 **70**: 62–79.

948 R Core Team. 2023. R: A Language and Environment for Statistical Computing. R Foundation for
949 Statistical Computing.

950 Radersma, R., Tinbergen, J.M. & Komdeur, J. 2011. Do brood sex ratio, nestling development
951 and sex affect fledging timing and order? An experimental study on great tits. *Animal*
952 *Behaviour* **81**: 69–75.

953 Rodríguez, S., van Noordwijk, A.J., Álvarez, E. & Barba, E. 2016. A recipe for postfledging survival
954 in great tits *Parus major*: be large and be early (but not too much). *Ecology and*
955 *Evolution* **6**: 4458–4467.

956 Searle, S.R. 1961. Phenotypic, Genetic and Environmental Correlations. *Biometrics* **17**: 474–480.
957 [Wiley, International Biometric Society].

958 Service, P.M. & Rose, M.R. 1985. Genetic covariation among life-history components: the effect
959 of novel environments. *Evolution* **39**: 943–945.

960 Siepielski, A.M., DiBattista, J.D. & Carlson, S.M. 2009. It's about time: the temporal dynamics of
961 phenotypic selection in the wild. *Ecology Letters* **12**: 1261–1276.

962 Simmonds, E.G., Cole, E.F., Sheldon, B.C. & Coulson, T. 2020. Testing the effect of quantitative
963 genetic inheritance in structured models on projections of population dynamics. *Oikos*
964 **129**: 559–571.

965 Sinervo, B. & Licht, P. 1991. Proximate Constraints on the Evolution of Egg Size, Number, and
966 Total Clutch Mass in Lizards. *Science* **252**: 1300–1302. American Association for the

967 Advancement of Science.

968 Smith, C.C. & Fretwell, S.D. 1974. The Optimal Balance between Size and Number of Offspring.
969 *The American Naturalist* **108**: 499–506. [The University of Chicago Press, The American
970 Society of Naturalists].

971 Smith, H.G., Källander, H. & Nilsson, J.-Å. 1989. The Trade-Off Between Offspring Number and
972 Quality in the Great Tit *Parus major*. *Journal of Animal Ecology* **58**: 383–401. [Wiley,
973 British Ecological Society].

974 Smith, H.G. & Nilsson, J.-Å. 1987. Intraspecific Variation in Migratory Pattern of a Partial
975 Migrant, the Blue Tit (*Parus caeruleus*): An Evaluation of Different Hypotheses. *The Auk*
976 **104**: 109–115.

977 Sogard, S.M. 1997. Size-Selective Mortality in the Juvenile Stage of Teleost Fishes: A Review.
978 *Bulletin of Marine Science* **60**: 1129–1157.

979 Taborsky, B., Skubic, E. & Bruintjes, R. 2007. Mothers adjust egg size to helper number in a
980 cooperatively breeding cichlid. *Behavioral Ecology* **18**: 652–657.

981 Tschirren, B., Fitze, P.S. & Richner, H. 2003. Sexual dimorphism in susceptibility to parasites and
982 cell-mediated immunity in great tit nestlings. *Journal of Animal Ecology* **72**: 839–845.

983 van de Pol, M., Bailey, L.D., McLean, N., Rijdsdijk, L., Lawson, C.R. & Brouwer, L. 2016. Identifying
984 the best climatic predictors in ecology and evolution. *Methods in Ecology and Evolution*
985 **7**: 1246–1257.

- 986 van Noordwijk, A.J. & de Jong, G. 1986. Acquisition and Allocation of Resources: Their Influence
987 on Variation in Life History Tactics. *The American Naturalist* **128**: 137–142. The
988 University of Chicago Press.
- 989 Van Noordwijk, A.J., McCleery, R.H. & Perrins, C.M. 1995. Selection for the Timing of Great Tit
990 Breeding in Relation to Caterpillar Growth and Temperature. *Journal of Animal Ecology*
991 **64**: 451–458. [Wiley, British Ecological Society].
- 992 Vanhove, J. 2021. Collinearity isn't a disease that needs curing. *Meta-Psychology* **5**.
- 993 Verhulst, S., Perrins, C.M. & Riddington, R. 1997. Natal Dispersal of great tits in a patchy
994 environment. *Ecology* **78**: 864–872.
- 995 Wilkin, T.A., Garant, D., Gosler, A.G. & Sheldon, B.C. 2006. Density effects on life-history traits in
996 a wild population of the great tit *Parus major*: analyses of long-term data with GIS
997 techniques. *Journal of Animal Ecology* **75**: 604–615.
- 998 Williams, G.C. 1966. Natural Selection, the Costs of Reproduction, and a Refinement of Lack's
999 Principle. *The American Naturalist* **100**: 687–690. The University of Chicago Press.
- 1000 Wolf, J.B. & Wade, M.J. 2001. On the assignment of fitness to parents and offspring: whose
1001 fitness is it and when does it matter? *Journal of Evolutionary Biology* **14**: 347–356.
- 1002 Wright, J., Bolstad, G.H., Araya-Ajoy, Y.G. & Dingemanse, N.J. 2019. Life-history evolution under
1003 fluctuating density-dependent selection and the adaptive alignment of pace-of-life
1004 syndromes. *Biological Reviews* **94**: 230–247.