

Supplementary Material 2: model checks and diagnostics

Offspring sex ratio increases with male reproductive success in the polygynous southern elephant seals

Hassen Allegue Christophe Guinet Samantha C. Patrick Cécile Ribout
Coraline Bichet Olivier Lepais Denis Réale

Contents

| | | |
|----------|---|-----------|
| 1 | The model | 2 |
| 2 | Prior predictive checks | 2 |
| 2.1 | Informative priors | 2 |
| 2.2 | Weakly informative priors | 3 |
| 3 | The fitted model | 5 |
| 3.1 | With the informative priors | 5 |
| 3.2 | With the weakly informative priors | 6 |
| 4 | Visualizing MCMC draws and diagnostics | 7 |
| 5 | Posterior predictive checks | 10 |

In this study, we proposed the hypothesis that offspring sex-ratio adjustment has evolved in southern elephant seals allowing males to produce more offspring of the sex that will provide a higher fitness. To test this hypothesis, we modeled the probability of producing a son as a function of the male relative reproductive success and the male relative body length (a proxy for age) which was fitted in a Bayesian framework with the R package `brms` (Bürkner, 2017).

The aim of this document is to provide all information on the model we used. We present the model equation (Section 1), the prior predictive checks (Section 2), the fitted model output (Section 3), the diagnostics for the MCMC sampling draws (Section 4), and the posterior predictive checks (Section 5). Overall, the various checks and diagnostics do not show any particular concerns with our model.

1 The model

$$y_i \sim \text{Binomial}(n_i, p_i)$$

$$\text{Logit}(p_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$$

y_i is the number of sons in n_i offspring produced by the breeding male i , p_i is the underlying (latent) probability of producing a son for the breeding male i . p_i is modeled as a linear regression at the latent scale using a *logit* link function. β_0 is the intercept, β_1 is the slope of the $\text{Logit}(p_i)$ in response to the relative reproductive success (x_1) of the breeding male i , β_2 is the slope of the $\text{Logit}(p_i)$ in response to the relative body length (x_2) of the breeding male i , and ϵ_i is the residual value.

2 Prior predictive checks

2.1 Informative priors

The informative priors were defined according to a theoretical study (Fawcett et al., 2007) and a meta-analysis (Booksmythe et al., 2017). Both studies found that the effect size of the quality/attractiveness of the males on the offspring sex ratio is always weak and positive when significantly different from zero. We therefore used a normal distribution for β_1 and β_2 with a slightly positive mean ($\mu = 0.1$) and a low standard deviation ($\sigma = 0.1$). The prior distribution of the intercept (β_0) followed a normal distribution with a mean of 0 and a larger standard deviation of 0.3:

$$\beta_0 \sim \text{normal}(0, 0.3)$$

$$\beta_1 \sim \text{normal}(0.1, 0.1)$$

$$\beta_2 \sim \text{normal}(0.1, 0.1)$$

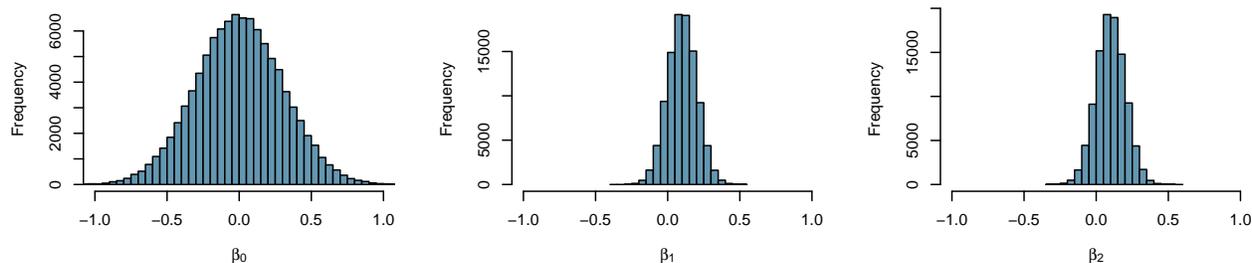


Figure S1: The histograms of the informative prior distributions of each of the model parameters: (β_0) the intercept, (β_1) the slope associated with the male relative reproductive success, and (β_2) the slope associated with the male relative body length.

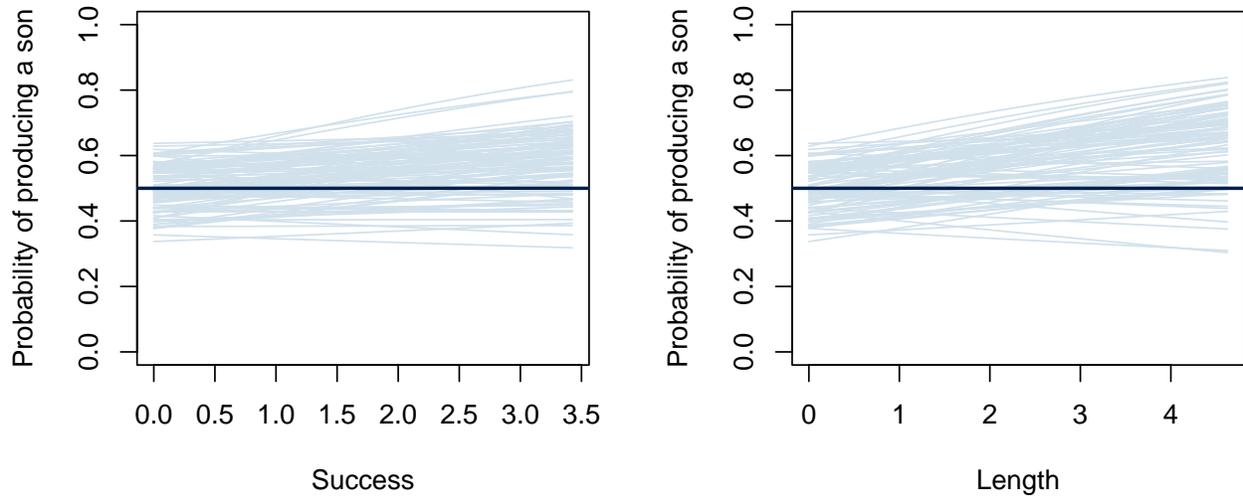


Figure S2: The probability of producing a son as a function of the male relative reproductive success (left) and the male relative body length (right) computed from the informative prior distributions. Each line represents a random sampling from the informative prior distributions of each of the model parameters. The darker horizontal line shows the probability of 0.5.

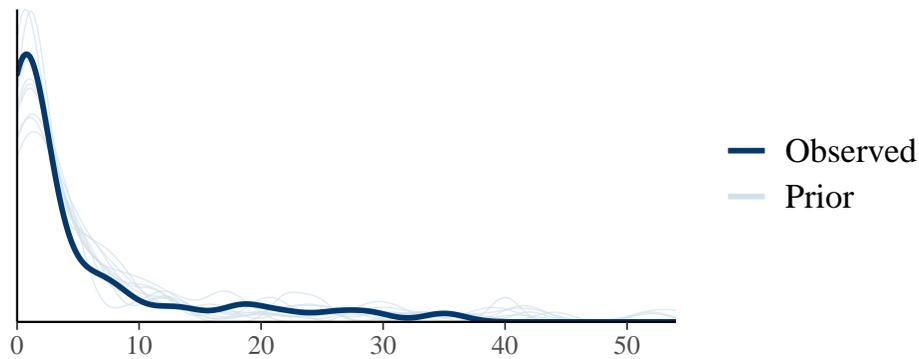


Figure S3: Density function of the observed data overlaid to 10 simulated informative prior predictive distributions.

2.2 Weakly informative priors

To evaluate the stability of our model, we also used weakly informative priors for the model parameters (i.e., β_0 , β_1 , and β_2) with normal distributions centered at 0 and with larger standard deviations:

$$\beta_0 \sim \text{normal}(0, 1)$$

$$\beta_1 \sim \text{normal}(0, 0.5)$$

$$\beta_2 \sim \text{normal}(0, 0.5)$$

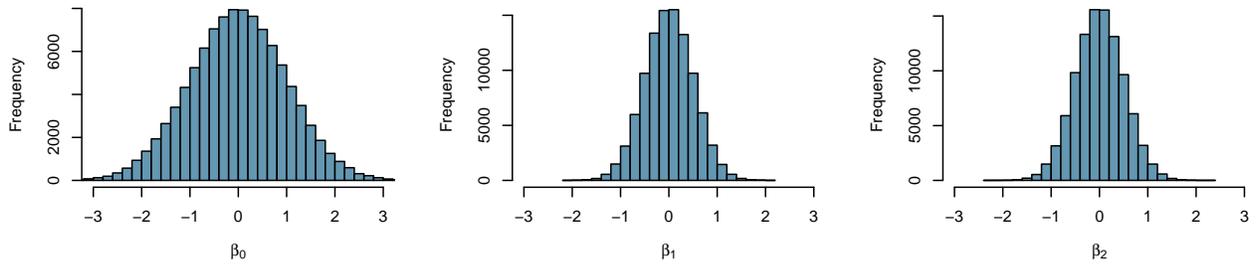


Figure S4: The histograms of the weakly informative prior distributions of each of the model parameters: (β_0) the intercept, (β_1) the slope associated with the male relative reproductive success, and (β_2) the slope associated with the male relative body length.

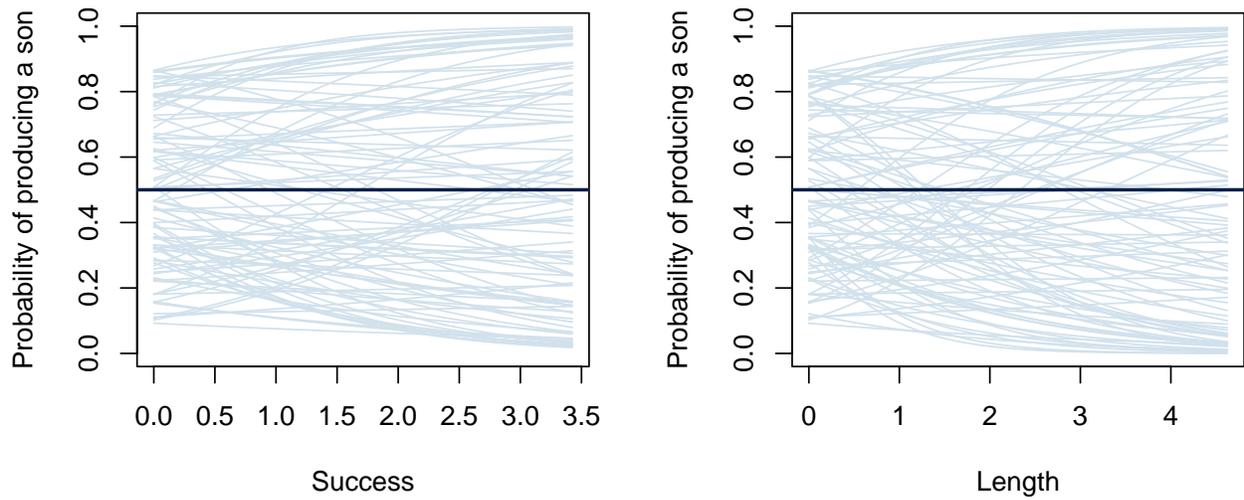


Figure S5: The probability of producing a son as a function of the male relative reproductive success (left) and the male relative body length (right) computed from the weakly informative prior distributions. Each line represents a random sampling from the weakly informative prior distributions of each of the model parameters. The darker horizontal line shows the probability of 0.5.

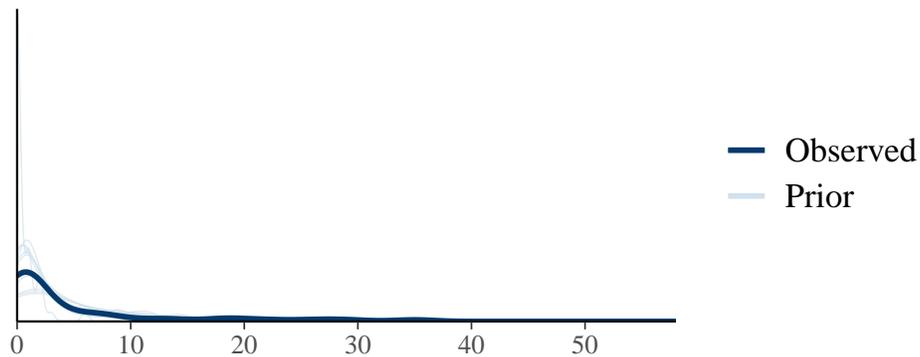


Figure S6: Density function of the observed data overlaid to 10 simulated weakly informative prior predictive distributions.

3 The fitted model

We fitted the model with the informative and the weakly informative priors.

3.1 With the informative priors

3.1.1 brms output summary

No divergent transitions was reported when fitting the model with the function `brm()`.

```
## Family: binomial
## Links: mu = logit
## Formula: Male | trials(N_pups) + weights(Weights, scale = TRUE) ~ 0 + Intercept + Success + Length
## Data: dat (Number of observations: 50)
## Samples: 4 chains, each with iter = 20000; warmup = 10000; thin = 4;
##           total post-warmup samples = 10000
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept    -0.21     0.18   -0.56    0.15 1.00    9542    9038
## Success       0.11     0.05    0.00    0.21 1.00    9820    9714
## Length        0.00     0.06   -0.11    0.12 1.00    9678    9395
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

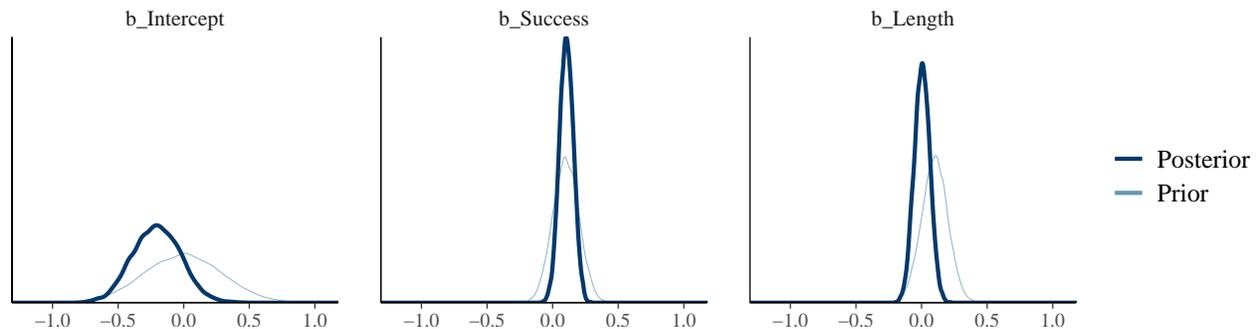


Figure S7: Density function of the posterior distributions overlaid to the informative prior distributions of each of the model parameters: (`b_Intercept`) the intercept, (`b_Success`) the slope associated with the male relative reproductive success, and (`b_Length`) the slope associated with the relative body length.

3.1.2 Testing the hypotheses

We use the function `hypothesis()` from the `brms` R package to test the hypothesis that $\beta_1 > 0$ against the null hypothesis that $\beta_1 = 0$ and the hypothesis that $\beta_2 > 0$ against the null hypothesis that $\beta_2 = 0$. The results show that $\beta_1 > 0$ is 41.2 times more likely than $\beta_1 = 0$ and that $\beta_2 > 0$ is 1.1 times more likely than $\beta_2 = 0$. Here, We conclude that we have weak evidence that $\beta_1 > 0$ and no evidence that $\beta_2 > 0$.

```
## Hypothesis Tests for class b:
## Hypothesis Estimate Est.Error CI.Lower CI.Upper Evid.Ratio Post.Prob Star
```

```

## 1 (Success) > 0      0.11      0.05      0.02      0.19      41.19      0.98      *
## ---
## 'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.
## '*': For one-sided hypotheses, the posterior probability exceeds 95%;
## for two-sided hypotheses, the value tested against lies outside the 95%-CI.
## Posterior probabilities of point hypotheses assume equal prior probabilities.

## Hypothesis Tests for class b:
##   Hypothesis Estimate Est.Error CI.Lower CI.Upper Evid.Ratio Post.Prob Star
## 1 (Length) > 0      0      0.06     -0.09      0.1      1.12      0.53
## ---
## 'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.
## '*': For one-sided hypotheses, the posterior probability exceeds 95%;
## for two-sided hypotheses, the value tested against lies outside the 95%-CI.
## Posterior probabilities of point hypotheses assume equal prior probabilities.

```

3.2 With the weakly informative priors

3.2.1 brms output summary

No divergent transitions was reported when fitting the model with the function `brm()`.

```

## Family: binomial
## Links: mu = logit
## Formula: Male | trials(N_pups) + weights(Weights, scale = TRUE) ~ 0 + Intercept + Success + Length
## Data: dat (Number of observations: 50)
## Samples: 4 chains, each with iter = 20000; warmup = 10000; thin = 4;
##           total post-warmup samples = 10000
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept    -0.15     0.29   -0.70    0.41 1.00     8820     9094
## Success       0.13     0.06   -0.00    0.25 1.00     8961     8152
## Length      -0.03     0.09   -0.22    0.15 1.00     8787     9371
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).

```

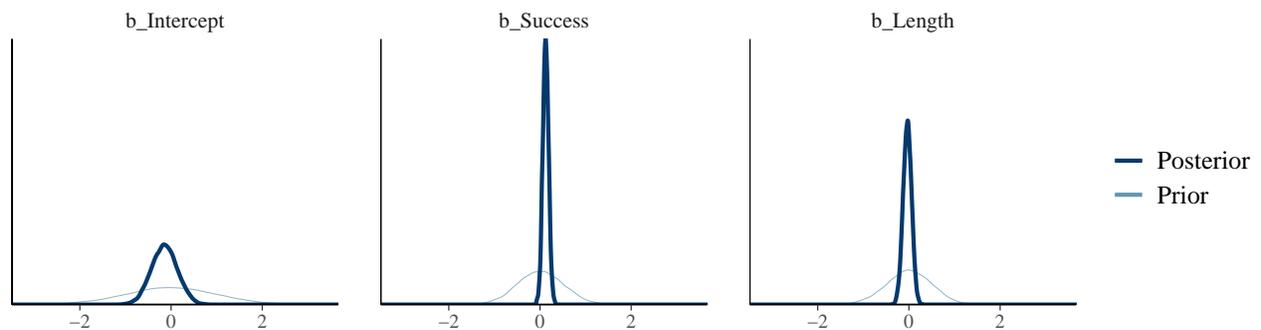


Figure S8: Density function of the posterior distributions overlaid to the weakly informative prior distributions of each of the model parameters: (`b_Intercept`) the intercept, (`b_Success`) the slope associated with the male relative reproductive success, and (`b_Length`) the slope associated with the relative body length.

3.2.2 Testing the hypotheses

The results show that $\beta_1 > 0$ is 37.5 times more likely than $\beta_1 = 0$ and that $\beta_2 > 0$ is 0.6 times more likely than $\beta_2 = 0$. We can see that we obtain the same results when fitting the model with the weakly informative priors compared to when we fit the model with the informative priors. We thus keep the model fitted with the informative priors as our final model and present the model diagnostics and checks in the next sections.

```
## Hypothesis Tests for class b:
## Hypothesis Estimate Est.Error CI.Lower CI.Upper Evid.Ratio Post.Prob Star
## 1 (Success) > 0 0.13 0.06 0.02 0.23 37.46 0.97 *
## ---
## 'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.
## '*': For one-sided hypotheses, the posterior probability exceeds 95%;
## for two-sided hypotheses, the value tested against lies outside the 95%-CI.
## Posterior probabilities of point hypotheses assume equal prior probabilities.

## Hypothesis Tests for class b:
## Hypothesis Estimate Est.Error CI.Lower CI.Upper Evid.Ratio Post.Prob Star
## 1 (Length) > 0 -0.03 0.09 -0.19 0.12 0.56 0.36
## ---
## 'CI': 90%-CI for one-sided and 95%-CI for two-sided hypotheses.
## '*': For one-sided hypotheses, the posterior probability exceeds 95%;
## for two-sided hypotheses, the value tested against lies outside the 95%-CI.
## Posterior probabilities of point hypotheses assume equal prior probabilities.
```

4 Visualizing MCMC draws and diagnostics

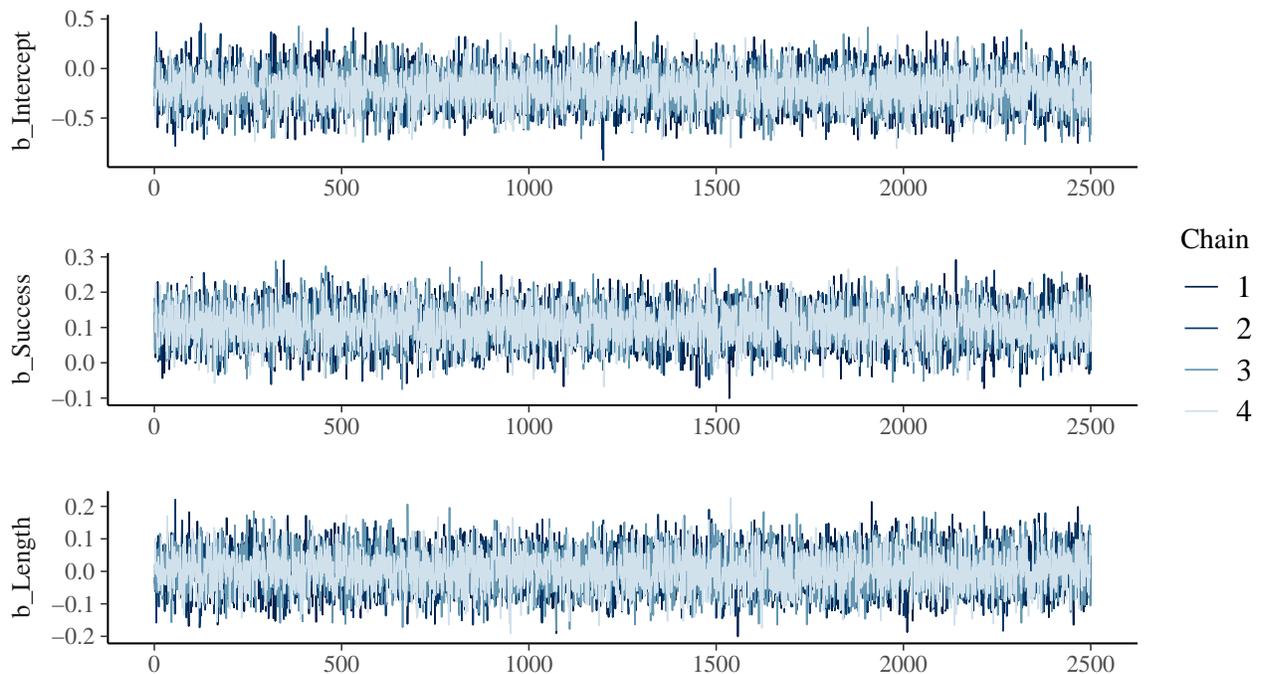


Figure S9: Time series (trace plot) of each of the Markov chains for each of the model parameters: (`b_Intercept`) the intercept, (`b_Success`) the slope associated with the male relative reproductive success, and (`b_Length`) the slope associated with the relative body length.

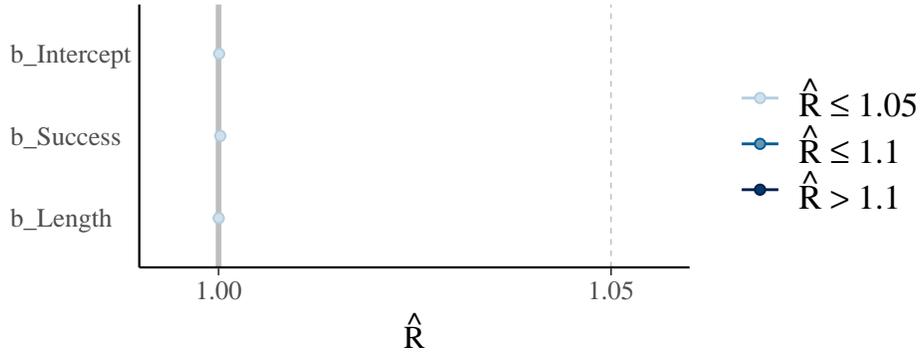


Figure S10: The potential scale reduction statistic ($split - \hat{R}$) which measures the ratio of the variance of draws within each chain to the variance of all draws across chains (Gelman et al., 2013) for each of the model parameters: ($b_Intercept$) the intercept, ($b_Success$) the slope associated with the male relative reproductive success, and (b_Length) the slope associated with the relative body length. If all chains converged similarly, these will be the same and \hat{R} will be one. If the chains have not converged to a common distribution, the \hat{R} statistic will be greater than one.

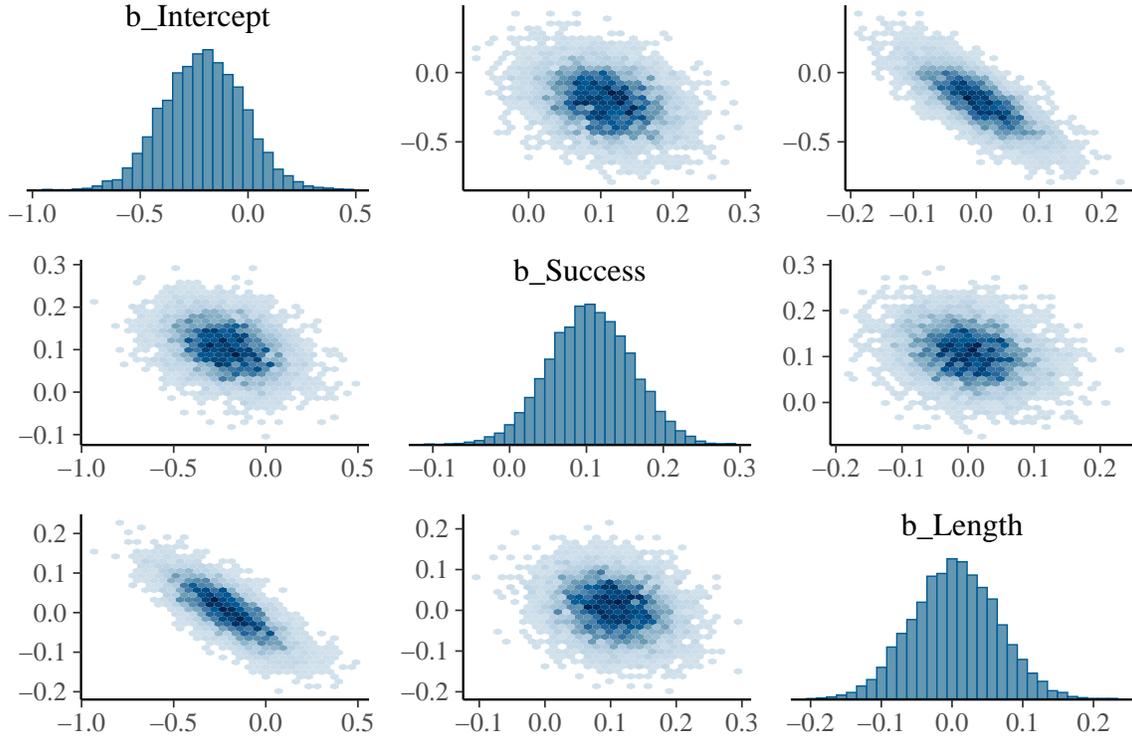


Figure S11: Histograms of the marginal posterior distribution (along the diagonal) of each of the model parameters: ($b_Intercept$) the intercept, ($b_Success$) the slope associated with the male relative reproductive success, and (b_Length) the slope associated with the relative body length. Bivariate plots are displayed as hex plots above (2 Markov chains) and below (2 other Markov chains) the diagonal.

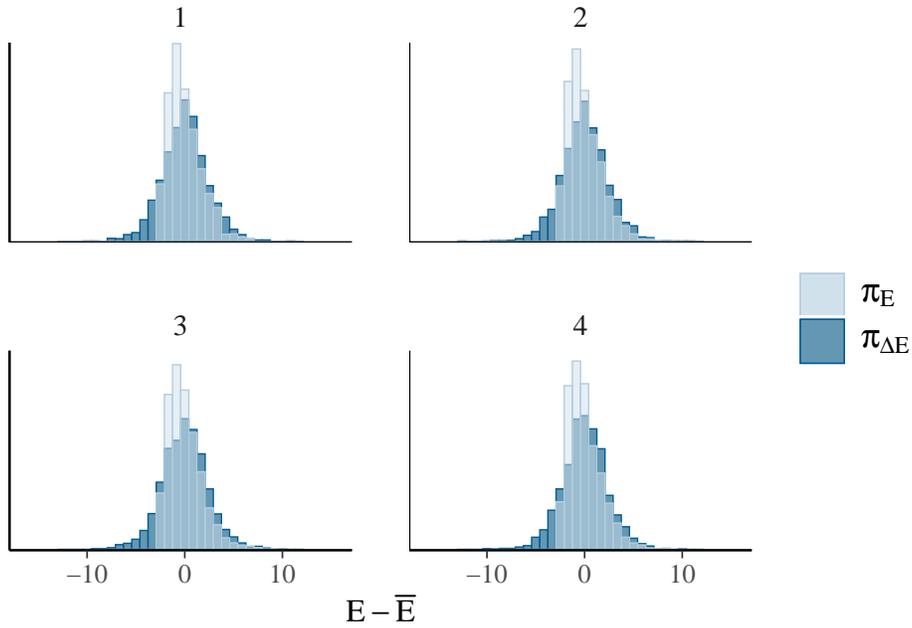


Figure S12: Overlaid histograms of the (centered) marginal energy distribution π_E and the first-differenced distribution $\pi_{\Delta E}$ for each of the Markov chains. Ideally both histograms should look the same (Betancourt, 2018).

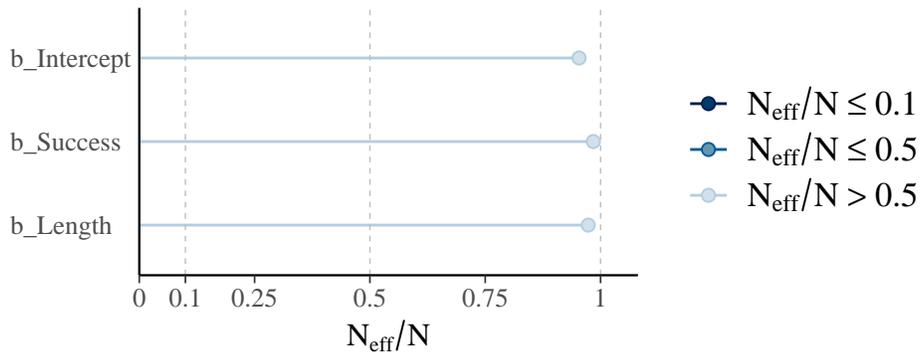


Figure S13: The ratio between the effective sample size (N_{eff}) to the total sample size (N) for each of the model parameters: ($b_Intercept$) the intercept, ($b_Success$) the slope associated with the male relative reproductive success, and (b_Length) the slope associated with the relative body length. The effective sample size is an estimate of the number of independent draws from the posterior distribution of the estimand of interest. The larger the ratio of N_{eff} to N the better (Gelman et al., 2013).

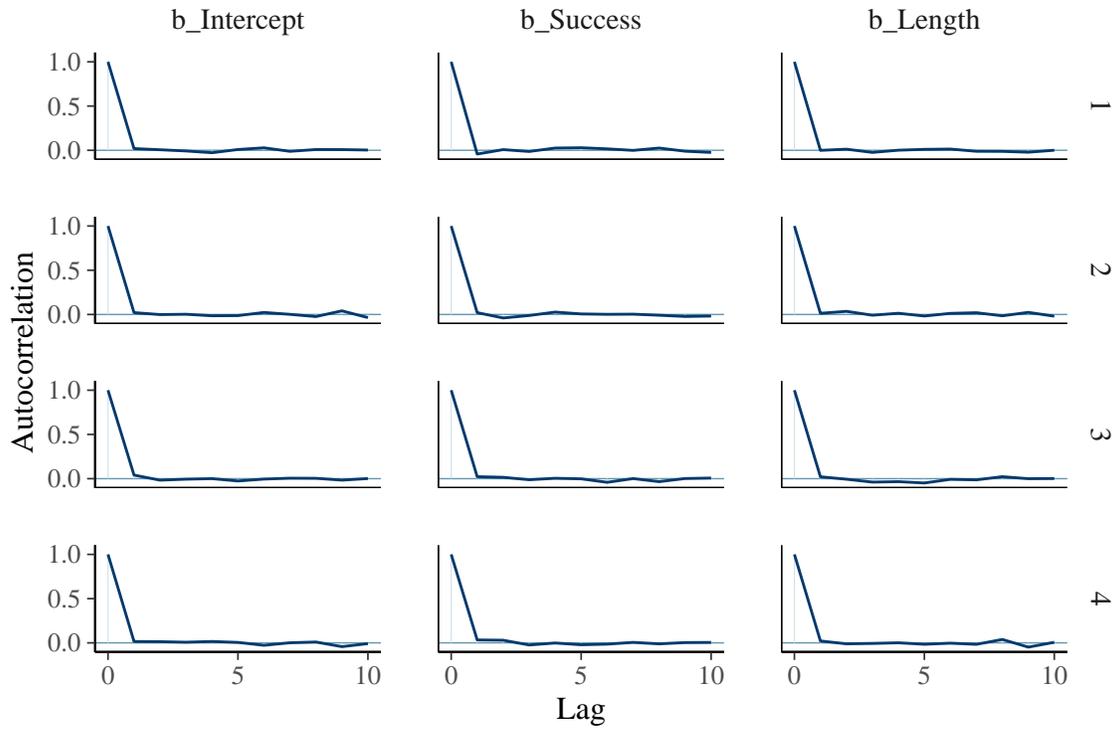


Figure S14: The autocorrelation (up to a lag of 10) of each of the Markov chains for each of the model parameters: (b_Intercept) the intercept, (b_Success) the slope associated with the male relative reproductive success, and (b_Length) the slope associated with the relative body length.

5 Posterior predictive checks

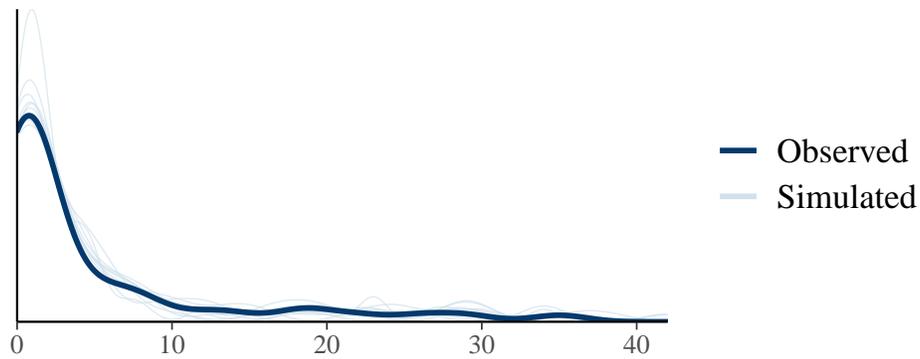


Figure S15: Density function of the observed data overlaid to 10 simulated posterior predictive distributions.

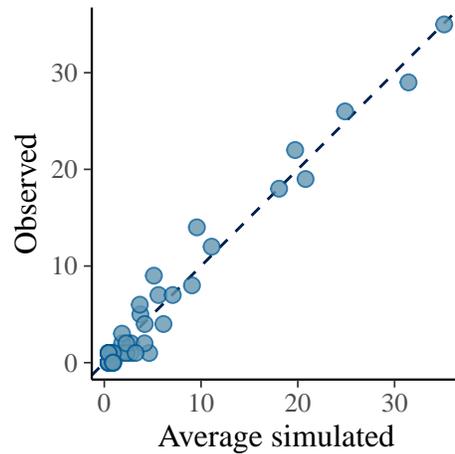


Figure S16: Scatter plot of the observed data as a function of the average of simulated posterior predictive distributions.

References

- Betancourt, M. (2018). A Conceptual Introduction to Hamiltonian Monte Carlo.
- Booksmythe, I., Mautz, B., Davis, J., Nakagawa, S., and Jennions, M. D. (2017). Facultative adjustment of the offspring sex ratio and male attractiveness: a systematic review and meta-analysis. *Biol. Rev.*, 92(1):108–134.
- Bürkner, P.-C. (2017). brms : an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.*, 80(1):1–28.
- Fawcett, T. W., Kuijper, B., Pen, I., and Weissing, F. J. (2007). Should attractive males have more sons? *Behav. Ecol.*, 18(1):71–80.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC Press., Boca Raton, FL, third edit edition.