

Title: How macroecology effects macroevolution and the interplay
2 between extinction intensity and selectivity in brachiopod
traxonomic survival.

4 **Running title:** Variation in the effects of traits on taxonomic survival

5 **Author:** Peter D Smits, psmits@uchicago.edu, Committee on Evolutionary
6 Biology, University of Chicago, IL, USA.

Keywords: extinction, brachiopods, Paleozoic, epicontinental seas, Bayesian

8 **Word count:**

Table count: 2

10 **Figure count:** 6

11 **Data archival location:** If accepted, all data and code necessary to duplicate
12 this analysis will be made available on DRYAD.

Abstract

As extinction intensity increases, how do the effects of traits on taxonomic survival change? Does the selective importance of certain traits increase while that of others decrease? Using a hierarchical Bayesian approach, I develop a model of how the effects of biological traits can vary with respect to extinction intensity, time of origination, and in relation to each other. I analyze patterns of Paleozoic brachiopod genus durations and their relationship to geographic range, affinity for epicontinental seas versus open ocean environments, and body size. Additionally, I estimate the effects of being environmental generalists versus specialist on taxonomic survival by allowing environmental preference to have a nonlinear effect on duration. My analysis framework eschews the traditional distinction between background and mass extinction, and instead the entire time period is analyzed as parts of the same continuum. I find that the cohort-specific estimates of effects of geographic range and environmental preference are both negatively correlated with estimates of baseline extinction intensity. These results indicate that as extinction intensity increases, it is expected that both effects will decrease in magnitude but that change in effect of environmental preference will be greater than the change in the effect of geographic range. Additionally, I find support for greater survival in environmental generalists over specialists in all origination cohorts analyzed, though taxa favoring epicontinental environments are expected to have a slightly greater duration than open-ocean favoring taxa. These results support the hypothesis that as extinction intensity increases, overall extinction selectivity decreases.

1 Introduction

- 40 How do biological traits affect extinction risk? Biological traits are defined here
as descriptors of a taxon's adaptive zone, which is the set of all biotic–biotic and
42 biotic–abiotic interactions that a taxon can experience (Simpson, 1944). In
effect, these are descriptors of a taxon's broad-sense ecology. Jablonski (1986)
44 observed that during a mass extinction event, the effects of biological traits on
taxonomic survival decreased in size. However, this pattern was not the case for
46 the effect of geographic range on survival (Jablonski, 1986). There are multiple
possible macroevolutionary mechanisms which may underlie the pattern
48 observed by Jablonski (1986): the effect of geographic range on predictive
survival remains constant and those of other biological traits decrease, the effect
50 of geographic range in predicting survival increases and those of other biological
traits stay constant, or the effects of all traits decrease by different degrees.
- 52 While Jablonski (1986) phrased his conclusions in terms of background versus
mass extinction, this observation can be analyzed in a continuous variation
54 framework as there is no obvious distinction in terms of extinction rate between
these two states (Payne and Finnegan, 2007, Simpson and Harnik, 2009, Wang,
56 2003). Additionally, the Jablonski (1986) scenario has strong model structure
requirements in order to test its proposed macroevolutionary mechanism; not
58 only do the taxon trait effects need to be modeled, but the correlation between
trait effects need to be modeled as well.
- 60 Conceptually, taxon survival can be considered an aspect of “taxon fitness”
along with expected lineage specific branching/origination rate (Cooper, 1984,
62 Palmer and Feldman, 2012). A taxon with a beneficial trait should persist for
longer, on average, than a taxon without that beneficial trait. Here I model
64 brachiopod taxon durations because trait based differences in extinction risk

should manifest as differences in taxon durations. Brachiopods are an ideal
group for this study as they are well known for having an exceptionally
complete fossil record (Foote, 2000b, Foote and Raup, 1996). I focus on the
brachiopod record from the post-Cambrian Paleozoic, from the start of the
Ordovician (approximately 485 My) through the end Permian (approximately
252 My) as this represents the time of greatest global brachiopod diversity
(Alroy, 2010b).

The analysis of taxon durations, or time from origination to extinction, falls
under the purview of survival analysis, a field of applied statistics commonly
used in health care (Klein and Moeschberger, 2003) but has a long history in
paleontology (Crampton et al., 2016, Simpson, 1944, 1953, Smits, 2015, Van
Valen, 1973, 1979). I adopt a hierarchical Bayesian survival modeling approach,
which represents both a conceptual and statistical unification of the
paleontological dynamic and cohort survival analytic approaches (Baumiller,
1993, Crampton et al., 2016, Ezard et al., 2012, Foote, 1988, Raup, 1975, 1978,
Simpson, 2006, Smits, 2015, Van Valen, 1973, 1979).

1.1 Factors affecting brachiopod survival

Geographic range is widely considered the most important taxon trait for
estimating differences in extinction risk at nearly all times relative to other
traits, with large geographic range associated with low extinction risk (Finnegan
et al., 2012, Harnik et al., 2012, Jablonski, 1986, 1987, 2008, Jablonski and Roy,
2003, Payne and Finnegan, 2007), though Foote and Miller (2013) find that this
generalization does not hold in the Mesozoic. For the Paleozoic, however, I
expect this to hold true for the entire period analyzed.

Epicontinental seas are a shallow-marine environment where the ocean has

90 spread over the surface of a continental shelf with a depth typically less than
100m. In contrast, open-ocean coastline environments have much greater
92 variance in depth, do not cover the continental shelf, and can persist during
periods of low sea level. Because of this, it is strongly expected that taxa which
94 favor epicontinental seas would be at great risk during periods of low sea levels,
such as during glacial periods, when epicontinental seas are drained. During the
96 Paleozoic (approximately 541–252 My), epicontinental seas were widely spread
globally but declined over the Mesozoic (approximately 252–66 My) and have
98 nearly disappeared during the Cenozoic (approximately 66–0 My) as open-ocean
coastlines became the dominant shallow-marine setting (Johnson, 1974, Miller
100 and Foote, 2009, Peters, 2008).

Miller and Foote (2009) demonstrated that during several mass extinctions taxa
102 associated with open-ocean environments tend to have a greater extinction risk
than those taxa associated with epicontinental seas. During periods of
104 background extinction, however, they found no consistent difference between
taxa favoring either environment. These two environment types represent the
106 primary environmental dichotomy observed in ancient marine systems (Miller
and Foote, 2009, Peters, 2008, Sheehan, 2001). Given these findings, I predict
108 that as extinction risk increases, the extinction risk associated open-ocean
environments should generally increase.

110 Because environmental preference is defined as the continuum between
occurring exclusively in open-ocean environments versus epicontinental
112 environments, intermediate values are considered “generalists” in the sense that
they favor neither end member. A long standing hypothesis is that generalists or
114 unspecialized taxa will have greater survival than specialists (Baumiller, 1993,
Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944). Because
116 of this, the effect of environmental preference was modeled as a quadratic

function where a concave down relationship between preference and expected
118 duration indicates that generalists are favored over specialists end-members.

Body size, measured as shell length, is also considered as a potential trait that
120 may influence extinction risk (Harnik, 2011, Payne et al., 2014). Body size is a
proxy for metabolic activity and other correlated life history traits (Payne et al.,
122 Harnik et al. (2014) analyzed the effect of body size selectivity in
Devonian brachiopods in both a phylogenetic and aphylogenetic context; finding
124 that body size was not found to be associated with differences in taxonomic
duration. It has also been found that, in the case of bivalves, for certain
126 subclades, body size can be as important a factor as geographic range size in
determining extinction risk (Harnik, 2011). Given these results, I expect that if
128 body size has any effect on brachiopod taxonomic survival it is very small.

It is well known that, given the incompleteness of the fossil record, the observed
130 duration of a taxon is an underestimate of that taxon's true duration (Alroy,
2014, Foote and Raup, 1996, Liow and Nichols, 2010, Solow and Smith, 1997,
132 Wagner and Marcot, 2013, Wang and Marshall, 2004). Because of this, the
concern is that a taxon's observed duration may reflect its relative chance of
134 being sampled and not any its relationships to the covariates of interest. In this
case, for sampling ability to be a confounding factor there must be consistent
136 relationship between sampling ability and duration (i.e. greater sampling, longer
duration or *vice versa*). If there is no relationship between sampling ability and
138 duration than interpretation can be made clearly; while observed durations are
obviously truncated true durations, a lack of a relationship would indicate that
140 the amount and form of this truncation is independent of the taxon's duration.

2 Materials and Methods

¹⁴² 2.1 Fossil occurrence information

The brachiopod dataset analyzed here was sourced from the Paleobiology
¹⁴⁴ Database (<http://www.paleodb.org>) which was then filtered based on taxonomic
(Rhynchonellata), temporal (post-Cambrian Paleozoic), stratigraphic, and other
¹⁴⁶ occurrence information used in this analysis. These filtering criteria are very
similar to those from Foote and Miller (2013) with an additional constraint of
¹⁴⁸ being present in the body size data set from Payne et al. (2014). Epicontinental
versus open-ocean assignments for each fossil occurrence are partially based on
¹⁵⁰ those from Miller and Foote (2009), with additional occurrences assigned
similarly (Miller and Foote, personal communication). Analyzed occurrences
¹⁵² were restricted to those with paleolatitude and paleolongitude coordinates,
assignment to either epicontinental or open-ocean environment, and belonging
¹⁵⁴ to a genus present in the body size dataset (Payne et al., 2014).

Fossil occurrences were analyzed at the genus level which is common for
¹⁵⁶ paleobiological, macroevolution, or macroecological studies of marine
invertebrates (Alroy, 2010b, Foote and Miller, 2013, Harnik et al., 2012,
¹⁵⁸ Kiessling and Aberhan, 2007, Miller and Foote, 2009, Nürnberg and Aberhan,
2013, 2015, Payne and Finnegan, 2007, Simpson and Harnik, 2009, Vilhena
¹⁶⁰ et al., 2013). While species diversity dynamics are frequently of much greater
interest than those of higher taxa (though see Foote 2014, Hoehn et al. 2015),
¹⁶² the nature of the fossil record makes accurate and precise taxonomic
assignments at the species level difficult for all occurrences. In particular, the
¹⁶⁴ simplicity of brachiopod external morphology and the quality of preservation
makes it very difficult non-experts to identify specimens to below the genus
¹⁶⁶ level. As such, the choice to analyze genera as opposed to species was in order

to assure a minimum level of confidence and accuracy in the data analyzed here.

168 Genus duration was calculated as the number of geologic stages from first
169 appearance to last appearance, inclusive. Durations were based on geologic
170 stages as opposed to millions of years because of the inherently discrete nature
171 of the fossil record; dates are not assigned to individual fossils themselves but
172 instead fossils are assigned to a geological interval which represents some
173 temporal range. In this analysis, stages are effectively irreducible temporal
174 intervals in which taxa may occur. Genera with a last occurrence in or after
175 Changhsingian stage were right censored at the Changhsingian; genera with a
176 duration of only one stage were left censored (Klein and Moeschberger, 2003).
177 The covariates used to model genus duration were geographic range size (r),
178 environmental preference (v, v^2), body size (m), and sampling (s).

Geographic range was calculated using an occupancy approach. First, all
180 occurrences were projected onto an equal-area cylindrical map projection. Each
181 occurrence was then assigned to one of the cells from a 70×34 regular raster
182 grid placed on the map. Each grid cell represents approximately $250,000 \text{ km}^2$.
The map projection and regular lattice were made using shape files from
184 <http://www.naturalearthdata.com/> and the **raster** package for R (Hijmans,
185 2015). For each stage, the total number of occupied grid cells was calculated.
186 Then, for each genus, the number of grid cells occupied by that genus was
187 calculated. Dividing the genus occupancy by the total occupancy gives the
188 relative occupancy of that genus. Mean relative genus occupancy was then
189 calculated as the mean of the per stage relative occupancies of that genus.
190 Sampling was calcualted as the average “gap ratio” of all stages in which a
191 taxon existed. The basis of the gap ratio are range-through taxa, which are taxa,
192 for any stage t , that are present in stages $t - 1$ and $t + 1$ (Alroy, 2010a, Foote,
2000a, Foote and Raup, 1996). A gap is when a range-through taxon is not

¹⁹⁴ present in stage t . The gap ratio for a stage is the number of gaps divided by the
¹⁹⁵ total number of range-through taxa, and the gap ratio for an individual taxon is
¹⁹⁶ the average of all gap ratios for each stage of the taxon's duration. Additional
¹⁹⁷ occurrence information for the late Cambrian and earliest Trassic were used to
¹⁹⁸ calculate the gap ratios of the first and last stages included in this analysis.

²⁰⁰ Environmental preference was defined as probability of observing the ratio of
²⁰¹ epicontinental occurrences to total occurrences ($\theta_i = e_i/E_i$) or greater given the
²⁰² background occurrence probability θ'_i as estimated from all other taxa occurring
at the same time (e'_i/E'_i). This measure of environmental preference is
expressed.

$$\begin{aligned} p(\theta'_i|e'_i, E'_i) &\propto \text{Beta}(e'_i, E'_i - e'_i)\text{Beta}(1, 1) \\ &= \text{Beta}(e'_i + 1, E'_i - e'_i + 1), \end{aligned} \tag{1}$$

²⁰⁴ where v is the percent of the distribution defined in equation 1 less than or
equal to θ_i .

²⁰⁶ Body size, measured as shell length, was sourced directly from Payne et al.
(2014).

²⁰⁸ Prior to analysis, geographic range was logit transformed and body size was
²⁰⁹ natural-log transformed. All covariates were then standardized by subtracting
²¹⁰ the mean from all values and dividing by twice its standard deviation, which
follows Gelman and Hill (2007). This standardization means that all regression
²¹² coefficients are comparable as the expected change per 1-unit change in the
rescaled covariates. D is defined as the total number of covariate effects
²¹⁴ included in the model.

2.2 Analytical approach

²¹⁶ Hierarchical modelling is a statistical approach which explicitly takes into
²¹⁸ account the structure of the observed data in order to model both the within
and between group variance (Gelman et al., 2013, Gelman and Hill, 2007). The
²²⁰ units of study (e.g. genera) each belong to a single group (e.g. origination
cohort). Each group is considered a draw from a shared probability distribution
of all cohorts, observed and unobserved. The group-level parameters are then
²²² estimated simultaneously as the other parameters of interest (e.g. covariate
effects) (Gelman et al., 2013). The subsequent estimates are partially pooled
²²⁴ together, where parameters from groups with large samples or effects remain
large while those of groups with small samples or effects are pulled towards the
²²⁶ overall group mean. All covariate effects (regression coefficients), as well as the
intercept term (baseline extinction risk), were allowed to vary by group
²²⁸ (origination cohort). The covariance between covariate effects was also modeled.

²³⁰ Genus durations were assumed to follow a Weibull distribution which allows for
age-dependent extinction (Klein and Moeschberger, 2003): $y \sim \text{Weibull}(\alpha, \sigma)$.
²³² The Weibull distribution has two parameters: scale σ , and shape α . When
 $\alpha = 1$, σ is equal to the expected duration of any taxon. α is a measure of the
²³⁴ effect of age on extinction risk where values greater than 1 indicate that
extinction risk increases with age, and values less than 1 indicate that
extinction risk decreases with age. Note that the Weibull distribution is
²³⁶ equivalent to the exponential distribution when $\alpha = 1$.

²³⁸ The scale parameter σ was modeled as a regression following Kleinbaum and
Klein (2005) with both varying intercept and varying slopes; this is expressed

$$\sigma_i = \exp\left(\frac{-\mathbf{X}_i B_{j[i]}}{\alpha_{j[i]}}\right) \quad (2)$$

where i indexes across all observations, $j[i]$ is the cohort membership of the i th
240 observation, X is a $N \times (D + 1)$ matrix of covariates along with a column of 1's
for the intercept term, and B is a $J \times D$ matrix of cohort-specific regression
242 coefficients.

Each of the rows of matrix B are modeled as realizations from a multivariate
244 normal distribution with length D location vector μ and $J \times J$ covariance
matrix Σ : $B_j \sim \text{MVN}(\mu, \Sigma)$. The covariance matrix was then decomposed into
246 a length J vector of scales τ and a $J \times J$ correlation matrix Ω , defined

$$\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau) \quad (3)$$

where “diag” indicates a diagonal matrix.

248 The elements of μ were given independent normally distributed priors. The
effects of geographic range size and the breadth of environmental preference
250 were given informative priors reflecting the previous findings while the others
were given weakly informative favoring that covariate having no effect. The
252 correlation matrix Ω was given an almost uniform LKJ distributed prior

following CITATION STAN manual. These priors are defined

$$\begin{aligned}
\mu_0 &\sim \mathcal{N}(0, 5) \\
\mu_r &\sim \mathcal{N}(-1, 1) \\
\mu_v &\sim \mathcal{N}(0, 1) \\
\mu_{v^2} &\sim \mathcal{N}(1, 1) \\
\mu_m &\sim \mathcal{N}(0, 1) \\
\mu_s &\sim \mathcal{N}(0, 1) \\
\tau &\sim \text{C}^+(1) \\
\Omega &\sim \text{LKJ}(2).
\end{aligned} \tag{4}$$

- ²⁵⁴ The shape parameter α was allowed to vary by cohort where
 $\log(\alpha_j) \sim \mathcal{N}(\alpha' + a_j, \sigma_\alpha)$. The effect of origination cohort a_j is modeled a
²⁵⁶ shared as normal distribution such that $a_j \sim \mathcal{N}(0, \sigma_a)$. α , σ_α and σ_a were given
weakly informative normal and half-Cauchy priors as appropriate:

$$\begin{aligned}
\alpha' &\sim \mathcal{N}(0, 1) \\
\sigma_a &\sim \text{C}^+(1) \\
\sigma_\alpha &\sim \text{C}^+(1).
\end{aligned} \tag{5}$$

- ²⁵⁸ The joint posterior was approximated using a Markov-chain Monte Carlo
routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn
²⁶⁰ Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic
programming language Stan (Stan Development Team, 2014). The posterior
²⁶² distribution was approximated from four parallel chains run for 10,000 steps
each, split half warm-up and half sampling and thinned to every 10th sample for
²⁶⁴ a total of 4000 posterior samples. Chain convergence was assessed via the scale

reduction factor \hat{R} where values close to 1 ($\hat{R} < 1.1$) indicate approximate
266 convergence. Convergence means that the chains are approximately stationary
and the samples are well mixed (Gelman et al., 2013).

268 The fit of the above model (the “full” model) was compared to the fits of three
other sub-models: constant α across cohorts, no sampling as a covariate, or both
270 constant α and no sampling covariate. These models were compared for
predicted out-of-sample predictive accuracy using both the widely-applicable
272 information criterion (WAIC) and leave-one-out cross-validation estimated via
Pareto-smoothed importance sampling (PSIS-LOO) (Vehtari and Gelman,
274 2015). Both of these are estimates of the out-of-sample predictive accuracy or
the expected quality of fit of the model to new data.

276 WAIC is a more fully Bayesian alternative to AIC or DIC (Gelman et al., 2013,
Watanabe, 2010); comparisons of WAIC values are useful for better
278 understanding the effect of model complexity on out-of-sample predictive
accuracy. The calculation of WAIC used here corresponds to the “WAIC 2”
280 formulation recommended by Gelman et al. (2013). Lower values of WAIC
indicate greater expected out-of-sample predictive accuracy than higher values.

282 PSIS-LOO is similar to WAIC in that it is an approximation of out-of-sample
predictive accuracy except its calculation is completely different (Vehtari and
284 Gelman, 2015, Vehtari et al., 2015b). Models comparison is done using a
leave-one-out crossvalidation information criterion (LOOIC), which is simply
286 the PSIS-LOO estimate multiplied by -2 so that it is on the deviance scale. As
with WAIC, models with lower values are expected to have a greater
288 out-of-sample predictive accuracy than models with greater values.

Calcuations of WAIC and PSIS-LOO for a model fit using Stan were done using
290 the R package “loo” (Vehtari et al., 2015a). See Vehtari et al. (2015b) for

detailed explanations of the calucations for both WAIC and PSIS-LOO.

292 Model adequacy was evaluated using a couple of posterior predictive checks.
Posterior predictive checks are a means for understanding model fit or
294 adequacy where the basic idea is that replicated data sets simulated from the
fitted model should be similar to the original data and systematic differences
296 between the simulations and observations indicate weaknesses of the model fit
(Gelman et al., 2013). For both approaches used here, 1000 posterior predictive
298 datasets were generated from 1000 unique draws from the posterior distribution
of each parameter. The two posterior predictive checks used in this analysis are
300 a comparison of a non-parameteric estimate of the survival function $S(t)$ from
the empirical dataset to the non-parameteric estimates of $S(t)$ from the 1000
302 posterior predictive datasets, and comparison of the observed genus durations to
the average posterior predictive estimate of $\log(\sigma)$ (Eq. 2). The former is to see
304 if simulated data has a similar survival pattern to the observed, while the latter
is to see if the model systematically over- or under- estimates taxon survival.

306 3 Results

The model with greatest estimated out-of-sample predictive accuracy, as
308 estimated via both WAIC and LOOIC, has both constant α and no sampling
covariate (Table 1). The effects of either of these potential model aspects are
310 inferred to be too small to account for the additional model complexity
necessary to estimate them. Because of this result, only the results from this
312 model will be presented here.

Comparison of the posterior predictive estimates of $S(t)$ to the empirical
314 estimate reveal few obvious biases except for the case of estimates from the far
right tail of observed durations (Fig. 1). This result is reinforced by the

Table 1: Comparison of widely applicable information criterion (WAIC) and leave-one-out cross-validation information criterion (LOOIC) values for the four models in this analysis. The WAIC and LOOIC values of the “best” model are in bold.

| model | WAIC | LOOIC |
|---------------------------------|----------------|----------------|
| constant α | 9080.39 | 4823.13 |
| constant α , no sampling | 4676.84 | 4621.95 |
| no sampling | 8250.22 | 5563.97 |
| full model | 759172.46 | 109417.64 |

³¹⁶ additional posterior predictive comparison where most estimates are not
 systematically biased, except for a consistent under-estimate of $\log(\sigma)$ for older
³¹⁸ taxa (Fig. 2). The results of both posterior predictive checks indicate that, for
 the majority of observations, model fit is generally not biased towards over- or
³²⁰ under-estimates of duration.

The cohort-level estimate of the effect of geographic range size indicates that as
³²² a taxon’s geographic range increases, that taxon’s duration is expected to
 increase (Table 2). Given the estimates of μ_r and τ_r , there is a less than 0.008%
³²⁴ (± 0.05 SD) probability that this relationships would be reversed
 $(\Pr(\mathcal{N}(\mu_r, \tau_r) > 0))$. The between-cohort variance in the cohort-specific
³²⁶ estimates τ_r are the lowest of all the regression coefficients (Table 2).

Body size is estimated to have no effect on taxon duration, with the estimate
³²⁸ being nearly 0 (Table 2). The variance between the cohort-specific estimates of
 the effect of body size τ_m is estimated to be greater than the variance of
³³⁰ between-cohort estimates of the effect of geographic range size τ_r .

The group-level estimate of the effect of environmental preference is estimated
³³² from both μ_v and μ_{v^2} .

The estimate of μ_v indicates that epicontinental favoring taxa are expected to
³³⁴ have a greater duration than open-ocean favoring taxa (Table 1). Additionally,

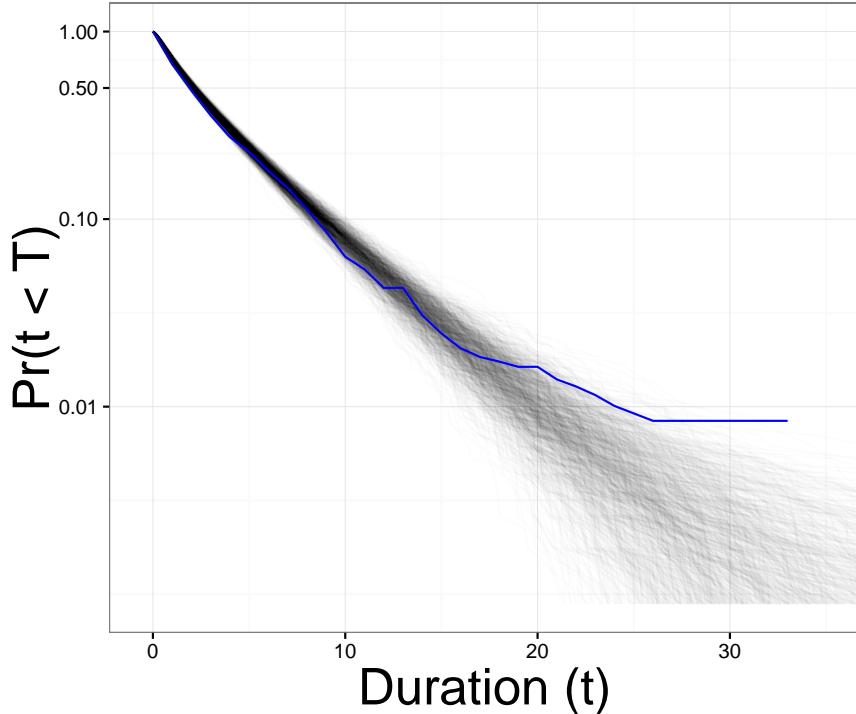


Figure 1: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 1000 posterior predictive data sets (black). $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T . The vertical axis is log10 transformed.

given the estimate of between-cohort variance τ_v , there is approximately 19%
336 (± 8 SD) probability that, for any given cohort, taxa favoring open-ocean
 environments would have a greater expected duration than taxa favoring
338 epicontinental environments ($\Pr(\mathcal{N}(\mu_v, \tau_v) > 0)$).

The estimate of μ_{v^2} indicates that the overall relationship between
340 environmental preference and $\log(\sigma)$ is concave down (Fig. 3), with only a 1.9%
 (± 2.4 SD) probability that any given cohort is convex up ($\Pr(\mathcal{N}(\mu_{v^2}, \tau_{v^2}) < 0)$).
342 As above, notice the direction of the inequality given the negative sign in
 equation 2.

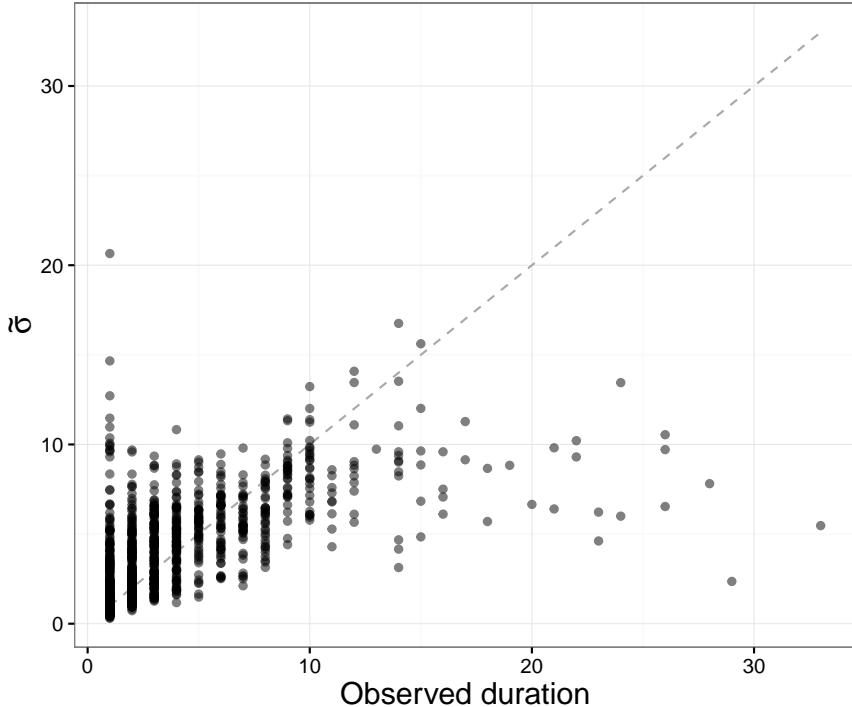


Figure 2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of $\log(\sigma)$. The dashed, diagonal line corresponds to $x = y$.

344 The cohort-specific estimates of all the regression coefficients demonstrate a lot
 345 of cohort to cohort variance, with no obvious trends. As indicated in Table 2
 346 and detectable visually (Fig. 4), the between-cohort estimates for β_0 , β_r , and
 347 β_m all have much lower variance than the between-cohort estimates of both β_v
 348 and β_{v^2} .

While most cohort-specific estimates are very similar to the overall cohort-level
 350 estimate, there are a few notable excursions away from the overall mean (Fig.
 351 4). There are simultaneous excursions in both β_0 and β_v for cohorts originating
 352 in the Givetian (387-382 My) and Frasnian (382-372 My) stages; both of which
 353 directly precede the end-Devonian mass extinction event at the

Table 2: Group-level estimates of the effects of biological traits on brachiopod generic survival. μ values are the location parameters of the effects, while τ values are the scale terms describing the variation between cohorts. The mean, standard deviation, 10th, 50th, and 90th quantiles of the posterior are presented.

| parameter | mean | standard deviation | 10% | 50% | 90% |
|--------------|-------|--------------------|-------|-------|-------|
| μ_i | -2.32 | 0.14 | -2.50 | -2.32 | -2.15 |
| μ_r | -0.76 | 0.11 | -0.91 | -0.76 | -0.62 |
| μ_v | -0.66 | 0.17 | -0.88 | -0.66 | -0.43 |
| μ_{v^2} | 2.88 | 0.31 | 2.48 | 2.88 | 3.27 |
| μ_m | 0.04 | 0.12 | -0.12 | 0.04 | 0.19 |
| τ_i | 0.50 | 0.10 | 0.37 | 0.49 | 0.63 |
| τ_r | 0.27 | 0.13 | 0.11 | 0.26 | 0.45 |
| τ_v | 0.76 | 0.16 | 0.56 | 0.74 | 0.97 |
| τ_{v^2} | 1.24 | 0.33 | 0.84 | 1.21 | 1.67 |
| τ_m | 0.47 | 0.12 | 0.33 | 0.47 | 0.63 |

³⁵⁴ Frasnian/Famennian boundary. These cohorts are marked by both a high extinction intensity and an increase in expected duration for taxa favoring ³⁵⁶ epicontinental environments over open-ocean ones.

Cohorts originating from the Silurian through the Early Devonian have a ³⁵⁸ slightly lower extinction intensity than the overall mean; these are from the Llandovery (443-443 My) through the Emsian (407-393 My). This is also a time ³⁶⁰ period where there is the lowest probability that epicontinental favoring taxa expected to have greater duration than open-ocean favoring taxa. Both the ³⁶² Silurian and Devonian periods are notable for having been periods with a mostly “hothouse” climate, with no polar icecaps and a high sea-level (Edwards ³⁶⁴ and Fanning, 1985, Joachimski et al., 2009, Munnecke et al., 2010).

The cohort-specific relationships between environmental preference and $\log(\sigma)$ ³⁶⁶ were calculated from the estimates of β_0 , β_v , and β_{v^2} (Fig. 5) and reflect how ³⁶⁸ these three parameters act in concert and not just individually (Fig. 4). Beyond results already discussed above in the context of the parameters individually, it is notable that the cohort originating in the Kungurian (279-272 My) is least like

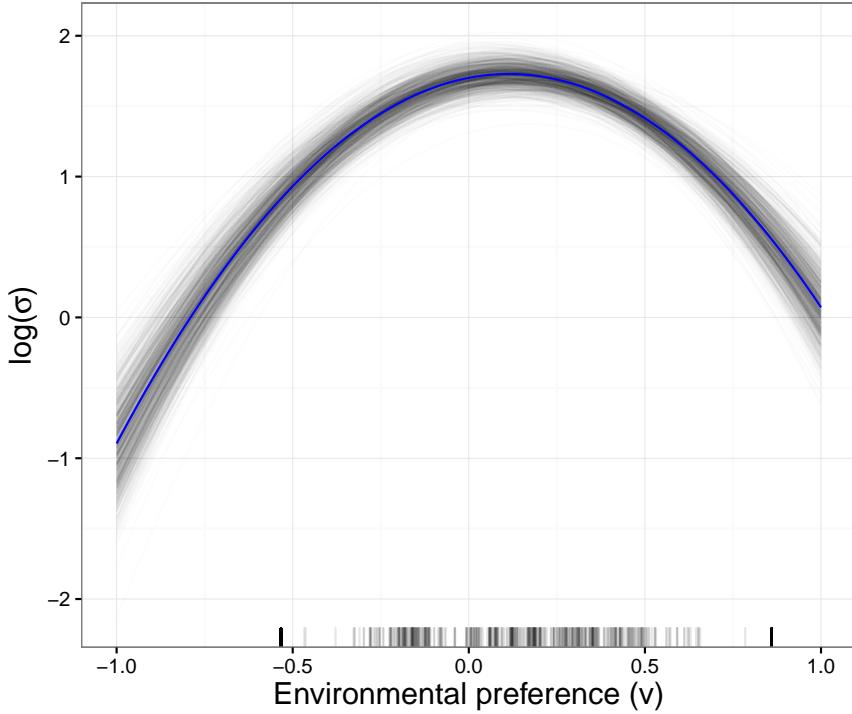


Figure 3: The overall expected relationship between environmental affinity v_i and a $\log(\sigma)$ when $r = 0$ and $m = 0$. Each grey line corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive distribution. The overall relationship is concave down with an optimum greater than 0, which means that taxa favoring epicontinental environments are expected to have longer durations than those favoring open-ocean environments.

³⁷⁰ the overall expected relationship and has the most sharply curved appearance
³⁷¹ due to a high estimate β_{v^2} (Fig. 4). This cohort is expected to have the biggest
³⁷² difference in extinction risk between environmental generalists and specialists. The
³⁷³ cohorts originating during the Emsian (407-393 My) and Frasnian (382 - 372
³⁷⁴ My) are tied most second for sharpness of curvature. The least sharply curved
³⁷⁵ cohorts include those originating during Tremadocian (484-477 My), Hirnantian
³⁷⁶ (445-443 My), Llandovery (443-433 My), and Ludlow (427-423 My). Except for
³⁷⁷ the Tremadocian cohort, all of these cohorts originate during the Silurian

³⁷⁸ through the Early Devonian range identified earlier as having lower expected extinction intensity than what is expected from the group-level estimate.

³⁸⁰ The correlations of the cohort-specific estimates of the regression coefficients are estimated as the off-diagonal elements of the correlation matrix Ω . Only two of ³⁸² the elements of Ω are distinguishable from 0: the correlation of β_0 (extinction intensity) with both β_r and β_v (Fig. 6).

³⁸⁴ There is an approximate 86% probability that the cohort-specific estimates of baseline extinction intensity β_0 and the effect of geographic range β_r are ³⁸⁶ negatively correlated; this means that for cohorts experiencing a lower extinction intensity, the effect of geographic range is expected to increase as ³⁸⁸ well, and *vice versa*.

Similarly, there is an approximate 99.9% probability that the cohort-specific ³⁹⁰ estimates of β_0 and β_v are negatively correlated; this means that as extinction intensity increases it is expected that epicontinental taxa become less favored ³⁹² over open-ocean environments. However, there is only a 19% (± 8.1 SD) probability that, for any given cohort, open-ocean environments will be favored.

³⁹⁴ There is only an approximate 68% probability that β_r and β_v are positively correlated. This lack of cross-correlation may be due in part to the much higher ³⁹⁶ between-cohort variance of the effect of environmental preference τ_v than the very small between-cohort variance in the effect of geographic range τ_r (Table ³⁹⁸ 2). The effect of geographic range might simply not vary enough relative to the much noisier environmental preference.

400 **4 Discussion**

The generating observation behind this study was that, as a mass extinction
402 event, the only biological trait that was found to affect extinction risk was
geographic range while traits that had previously been beneficial had no effect
404 (Jablonski, 1986). This observation raises two linked questions: how does the
effect of geographic range change with changing extinction intensity, and how
406 does the effect of other biological traits change with changing extinction
intensity?

408 I find that as intensity increases, the effect of geographic range is expected to
decreases. I also find that as intensity increases, the effect of environmental
410 preference is expected to be smaller. There is no evidence for a correlation
between the effect of geographic range and environmental preference. However,
412 the between-cohort variance in effect of geographic range is much less than the
between-cohort variance of the effect of environmental preference. As stated
414 above, this may underlie the lack of correlation between these two effects.

Additionally, the lower between-cohort variance in the effect of geographic range
416 versus that higher between-cohort variance implies that for cohorts with a
greater than average extinction intensity, the difference in the effect geographic
418 range and the group-level effect of geographic range will be smaller than the
difference between the effect of environmental preference and the group-level
420 effect of environmental preference.

I find consistent support for the “survival of the unspecialized,” with respect for
422 environmental preference, as a time-invariant generalization of brachiopod
survival; taxa with intermediate environmental preferences were found to have
424 lower extinction risk than taxa specializing in either epicontinental or
open-ocean environments (Fig. 3), though the curvature of the relationship

426 varies from nearly flat to very peaked (Fig. 5). However, this relationship is not
symetric about 0 as taxa favoring epicontinental environments are expected to
428 have a greater duration than taxa favoring open-ocean environments. This
description of environmental preference only describes one major aspect of a
430 taxon's environmental context, with factors such as bathymetry and
temperature being further descriptors of a taxon's adaptive zone (Harnik, 2011,
432 Harnik et al., 2012, Heim and Peters, 2011, Nürnberg and Aberhan, 2013);
inclusion of these factors would potentially improve our understanding of the
434 "survival of the unspecialized" hypothesis (Simpson, 1944).

Hopkins et al. (2014), in their analysis of niche conservatism and substrate
436 lithological preference in marine invertebrates, found that brachiopods were
among the least "conservative" groups; taxa were found to easily change
438 preference on short time scales. While substrate preference is not the same as
environmental preference, a question arises: do "generalists" have a greater
440 duration because they are "true" environmental generalists from orgination (e.g.
can inhabit both environments) or have taxa with longer durations evolved in
442 preference over their duration in the face of changing envrionments (i.e. niche
evolution and adaptation)? The nature of the analysis done here assumes that
444 traits are constant over the duration of the taxon; future work should explore
how environmental preference changes over lineage duration.

446 In this study I used the average quality of the record for the duration of a taxon
as a measure of the effect of sampling with the goal of determining if sampling
448 is acting as a confounding variable where taxa with shorter or longer durations
are associated with differences in sampling regime. This covariate, however, is
450 not included in the WAIC or LOOIC best model (Table 1). I conclude then that
the effect of sampling on duration is at most small and is not acting as a
452 confounding factor.

An alternative approach for specifically modeling survival than the method used here is the Cormack-Jolly-Seber (CJS) model is an alternative approach for modeling survival that can take into account imperfect observation (Liow et al., 2008, Liow and Nichols, 2010, Royle and Dorazio, 2008, Tomiya, 2013). This model is a type of hidden Markov model with an absorbing state (i.e. extinction). In this model, survival is defined as the probability of surviving from time t to time $t + 1$. Additionally, the effect of preservation and sighting is estimated as probability of observing a taxon that is present; this can extend the duration of a taxon beyond its last occurrence. This approach is a fundamentally different from the method used in the above analysis; for example, instead of modeling persistence from time t to time $t + 1$, I model the distribution taxon durations (lifetime) which then decomposes into a function describing survival ($S(t)$) and a function describing instantaneous extinction risk (Klein and Moeschberger, 2003, Royle and Dorazio, 2008).

The use of genera as the unit of the study and how to exactly interpret the effects of the biological traits is a remaining concern. For example, if any of the traits analyzed here are associated with increases in speciation rates, this might increase the duration of genera through self-renewal (Raup, 1991, 1994), which would be an example of the difference in biological pattern between species and genera (Jablonski, 1987, 2007, 2008). This could lead to a trait appearing to decrease generic level extinction risk by increasing species level origination rate instead of decreasing species level extinction risk. However, given the nature of the fossil record and maintaining a minimum level of data consistency/quality, there is no simple solution to decreasing this uncertainty in the interpretations of how the biological traits studied at the genus-level may translate to the species-level.

The model used here could be improved through either increasing the number of

⁴⁸⁰ analyzed taxon traits, expanding the hierarchical structure of the model to
⁴⁸¹ include other major taxonomic groups of interest, and the inclusion of explicit
⁴⁸² phylogenetic relationships between the taxa in the model as an additional
⁴⁸³ hierarchical effect. An example taxon trait that may be of particular interest is
⁴⁸⁴ the affixing strategy or method of interaction with the substrate of the taxon,
⁴⁸⁵ which has been found to be related to brachiopod survival where, for
⁴⁸⁶ cosmopolitan taxa, taxa that are attached to the substrate are expected to have
a greater duration than those that are not (Alexander, 1977).

⁴⁸⁸ It is theoretically possible to expand this model to allow for comparisons within
and between major taxonomic groups. This approach would better constrain the
⁴⁹⁰ brachiopod estimates while also allowing for estimation of similarities and
differences in cross-taxonomic patterns. The major issue surrounding this
⁴⁹² particular expansion involves finding a similarly well sampled taxonomic group
that is present during the Paleozoic. Example groups include Crinoidea,
⁴⁹⁴ Ostracoda, and other members of the “Paleozoic fauna” (Sepkoski Jr., 1981).

Taxon traits like environmental preference or geographic range (Hunt et al.,
⁴⁹⁶ 2005, Jablonski, 1987) are most likely heritable, at least phylogenetically
(Housworth et al., 2004, Lynch, 1991). Without phylogenetic context, this
⁴⁹⁸ analysis assumes that differences in extinction risk between taxa are independent
of the shared evolutionary history of those taxa (Felsenstein, 1985). In contrast,
⁵⁰⁰ the origination cohorts only capture shared temporal context. For example, if
taxon duration is phylogenetically heritable, then closely related taxa may have
⁵⁰² more similar durations as well as more similar biological traits. Without taking
into account phylogenetic similarity the effects of these biological traits would
⁵⁰⁴ be inflated solely due to inheritance. The inclusion of phylogenetic context as an
additional individual-level hierarchical effect independent of origination cohort
⁵⁰⁶ would allow for determining how much of the observed variability is due to

shared evolutionary history versus shared temporal context versus actual
508 differences associated with biological traits (Harnik et al., 2014, Smits, 2015).

In summary, patterns of Paleozoic brachiopod survival were analyzed using a
510 fully Bayesian hierarchical survival modelling approach while also eschewing the
traditional separation between background and mass extinction. I find that
512 cohort extinction intensity is negatively correlated with both the cohort-specific
effects of geographic range and environmental preference. These results imply
514 that as extinction intensity increases, it is expected that both effects will
decrease in magnitude. However, the change in effect of environmental
516 preference will be greater than the change in the effect of geographic range.
Additionally, I find support for greater survival in environmental generalists
518 over specialists in all origination cohorts analyzed; this is consistent with the
long standing “survival of the unspecialized” hypothesis (Baumiller, 1993, Liow,
520 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944, 1953, Smits,
2015). I also find that taxa favoring epicontinental environments are expected to
522 have a slightly greater duration than open-ocean favoring taxa. The results of
this analysis support the hypothesis that as extinction intensity increases,
524 overall extinction selectivity decreases.

Acknowledgements

526 I would like to thank K. Angielczyk, M. Foote, P. D. Polly, R. Ree, and G.
Slater for helpful discussion during the conception of this study. I'd also like to
528 thank D. Bapst, N. Pierrehumbert and M. Villarosa Garcia for additional
comments. An additional thank you to A. Miller for the epicontinental versus
530 open-ocean assignments. Finally, thank you to the reviewers for their helpful
comments. This entire study would not have been possible without the

⁵³² Herculean effort of the many contributors to the Paleobiology Database. In particular, I would like to thank J. Alroy, M. Aberhan, D. Bottjer, M. Clapham, ⁵³⁴ F. Fürsich, N. Heim, A. Hendy, S. Holland, L. Ivany, W. Kiessling, B. Kröger, A. McGowan, T. Olszewski, P. Novack-Gottshall, M. Patzkowsky, M. Uhen, L. ⁵³⁶ Villier, and P. Wager. This work was supported by a NASA Exobiology grant (NNX10AQ446) to A. Miller and M. Foote. I declare no conflicts of interest.

⁵³⁸ This is Paleobiology Database publication XXX.

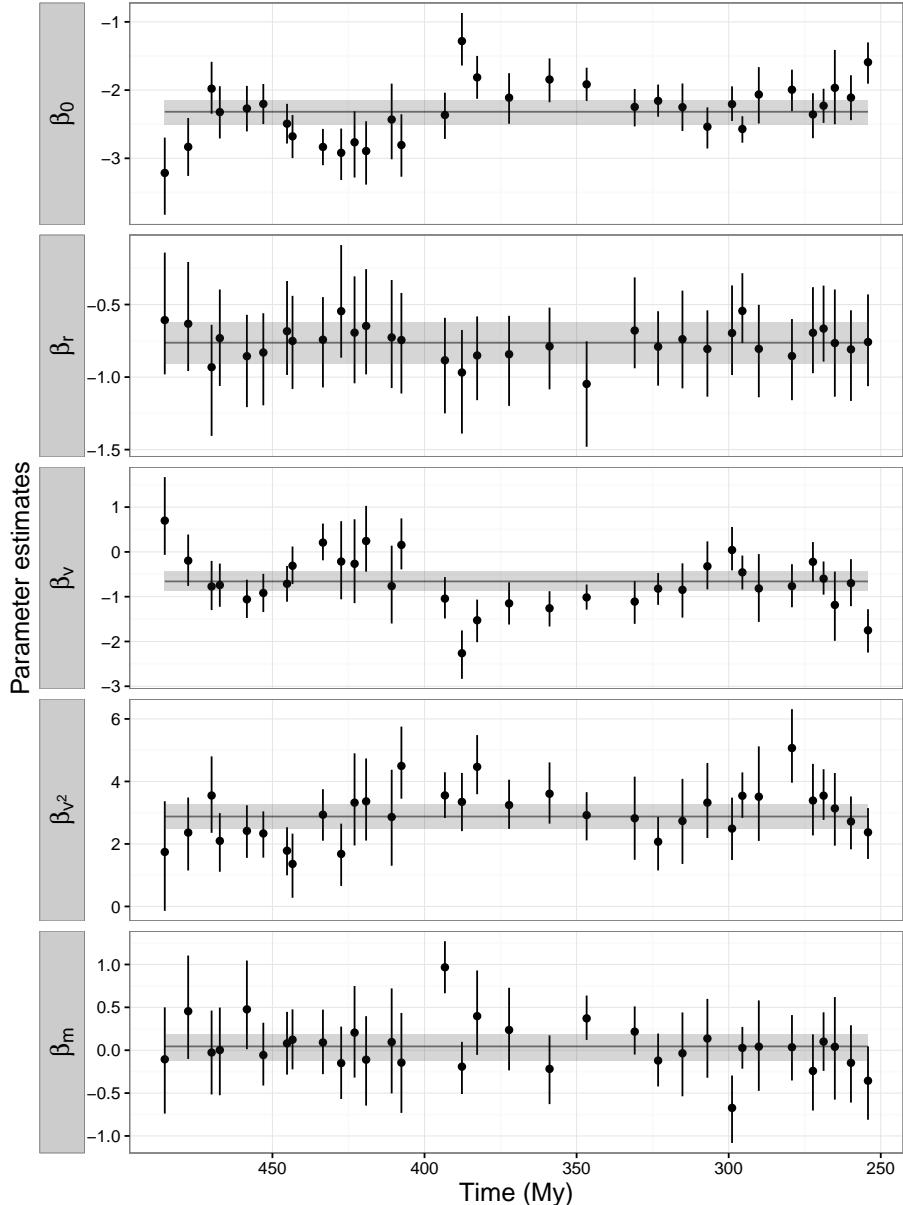


Figure 4: Comparison of cohort-specific estimates of β_0 , the effect of geographic range on extinction risk β_r , the effect of environmental preference β_v and β_{v^2} , and body size β_m . Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. Points are plotted at the midpoint of the cohorts stage of origination in millions of years before present (My). Black, horizontal lines are the overall estimates of covariate effects along with 80% credible intervals (shaded).

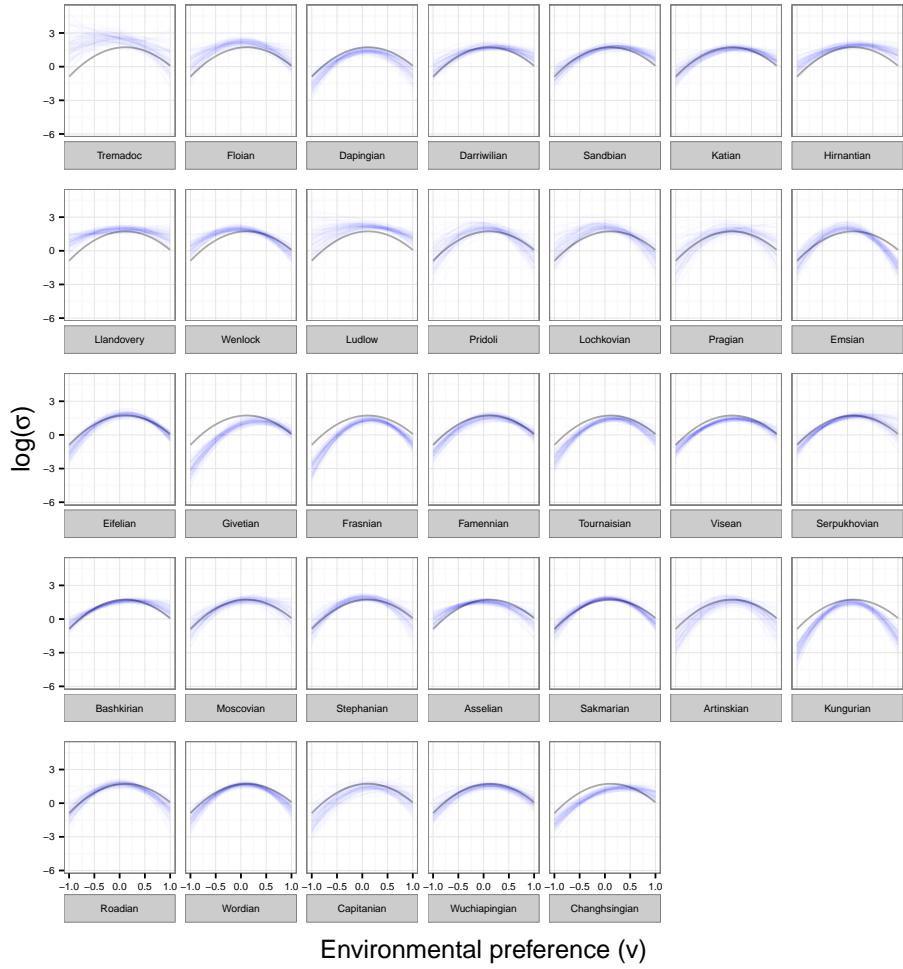


Figure 5: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on $\log(\sigma)$ to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the range of (transformed and rescaled) observed values of environmental preference. The oldest cohort is in the top-right and younger cohorts proceed left to right, with the youngest cohort being the right-most facet of the last row. Facet names correspond to the name of the stage in which that cohort originated.

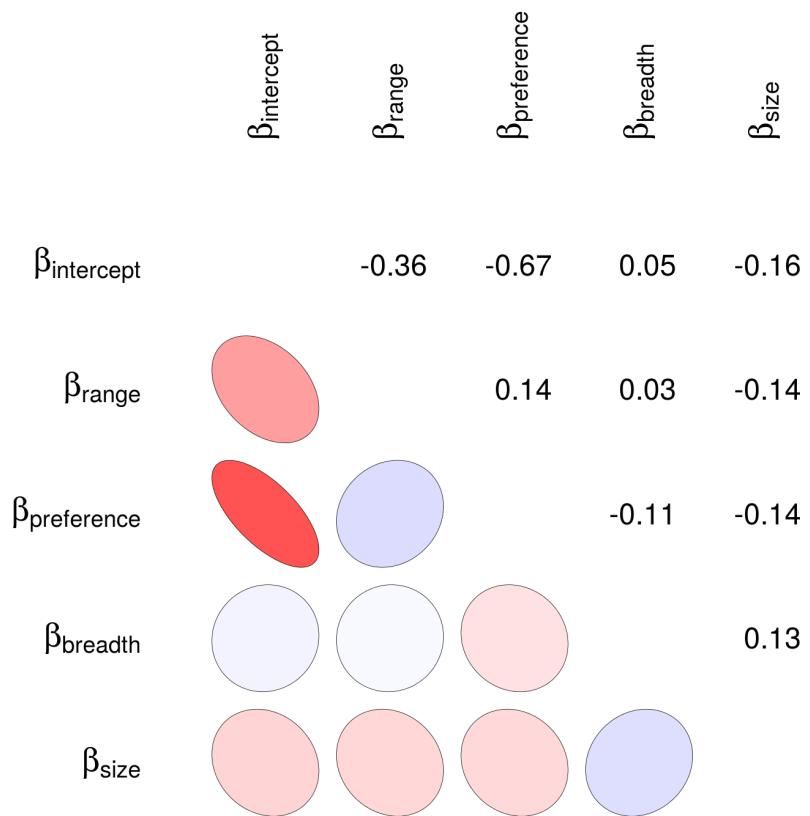


Figure 6: Mixed graphical and numerical representation of the correlation matrix Ω of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations are presented numerically (upper-triangle) and as idealized ellipses representing that much correlation (lower-triangle).

References

- 540 Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation
to the mode of stabilization on the substrate. *Palaeogeography,*
542 *Palaeoclimatology, Palaeoecology* 21:209–226.
- Alroy, J., 2010a. Fair sampling of taxonomic richness and unbiased estimation
544 of origination and extinction rates. Pp. 55–80, *in* J. Alroy and G. Hunt, eds.
Quantitative Methods in Paleobiology. The Paleontological Society.
- 546 ———, 2010b. The Shifting Balance of Diversity Among Major Marine Animal
Groups. *Science* 329:1191–1194.
- 548 ———, 2014. A simple Bayesian method of inferring extinction. *Paleobiology*
40:584–607.
- 550 Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of
filter morphology on evolutionary rates. *Paleobiology* 19:304–321.
- 552 Cooper, W. S., 1984. Expected time to extinction and the concept of
fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- 554 Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote, 2016.
Greenhouse/icehouse transition in the Late Ordovician marks a step change in
556 extinction regime in the marine plankton. *Proceedings of the National
Academy of Sciences P.* 201519092.
- 558 Edwards, D. and U. Fanning, 1985. Evolution and environment in the late
Silurian–early Devonian: the rise of pteridophytes. *Philosophical Transactions
560 of the Royal Society B: Biological Sciences* 309:147–165.
- Ezard, T. H. G., P. N. Pearson, T. Aze, and A. Purvis, 2012. The meaning of
562 birth and death (in macroevolutionary birth-death models). *Biology letters*
8:139–42.

- 564 Felsenstein, J., 1985. Phylogenies and the comparative method. *American
Naturalist* 125:1–15.
- 566 Finnegan, S., N. a. Heim, S. E. Peters, and W. W. Fischer, 2012. Climate
change and the selective signature of the Late Ordovician mass extinction.
568 *Proceedings of the National Academy of Sciences* 109:6829–6834.
- Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites.
570 *Paleobiology* 14:258–271.
- , 2000a. Origination and extinction components of taxonomic diversity:
572 general problems. *Paleobiology* 26:74–102.
- , 2000b. Origination and extinction components of taxonomic diversity:
574 Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- , 2014. Environmental controls on geographic range size in marine animal
576 genera. *Paleobiology* 40:440–458.
- Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine
578 animal genera. *Paleobiology* 39:171–192.
- Foote, M. and D. Raup, 1996. Fossil preservation and the stratigraphic ranges
580 of taxa. *Paleobiology* 22:121–140.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B.
582 Rubin, 2013. Bayesian data analysis. 3 ed. Chapman and Hall, Boca Raton,
FL.
- 584 Gelman, A. and J. Hill, 2007. Data Analysis using Regression and
Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- 586 Harnik, P. G., 2011. Direct and indirect effects of biological factors on

- extinction risk in fossil bivalves. *Proceedings of the National Academy of Sciences of the United States of America* 108:13594–9.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson, 2014.
- Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology* 40:675–692.
- Harnik, P. G., C. Simpson, and J. L. Payne, 2012. Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences* 279:4969–4976.
- Heim, N. A. and S. E. Peters, 2011. Regional environmental breadth predicts geographic range and longevity in fossil marine genera. *PloS one* 6:e18946.
- Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL <http://CRAN.R-project.org/package=raster>. R package version 2.3-24.
- Hoehn, K. B., P. G. Harnik, and V. L. Roth, 2015. A framework for detecting natural selection on traits above the species level. *Methods in Ecology and Evolution* .
- Hoffman, M. D. and A. Gelman, 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15:1351–1381.
- Hopkins, M. J., C. Simpson, and W. Kiessling, 2014. Differential niche dynamics among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- Housworth, E. A., P. Martins, and M. Lynch, 2004. The Phylogenetic Mixed Model. *The American Naturalist* 163:84–96.
- Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a comment on "On the heritability of geographic range sizes". *American Naturalist* 166:129–135.

- 612 Jablonski, D., 1986. Background and mass extincitons: the alternation of
macroevolutionary regimes. *Science* 231:129–133.
- 614 ———, 1987. Heritability at the species level: analysis of geographic ranges of
cretaceous mollusks. *Science* 238:360–363.
- 616 ———, 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- 618 ———, 2008. Species Selection: Theory and Data. *Annual Review of Ecology,
Evolution, and Systematics* 39:501–524.
- Jablonski, D. and K. Roy, 2003. Geographical range and speciation in fossil and
620 living molluscs. *Proceedings. Biological sciences / The Royal Society*
270:401–6.
- 622 Joachimski, M. M., S. Breisig, W. Buggisch, J. A. Talent, R. Mawson,
M. Gereke, J. R. Morrow, J. Day, and K. Weddige, 2009. Devonian climate
624 and reef evolution: Insights from oxygen isotopes in apatite. *Earth and
Planetary Science Letters* 284:599–609.
- 626 Johnson, J. G., 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine
628 benthic biodiversity dynamics through Triassic Jurassic time. *Paleobiology*
33:414–434.
- 630 Klein, J. P. and M. L. Moeschberger, 2003. *Survival Analysis: Techniques for
Censored and Truncated Data*. 2nd ed. Springer, New York.
- 632 Kleinbaum, D. G. and M. Klein, 2005. *Survival analysis: a self-learning text*. 2
ed. Springer, New York, NY.
- 634 Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively
unspecialized" using fossil crinoids. *The American naturalist* 164:431–43.

- 636 ———, 2007. Does versatility as measured by geographic range, bathymetric
range and morphological variability contribute to taxon longevity? *Global
Ecology and Biogeography* 16:117–128.
- 640 Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn,
and N. C. Stenseth, 2008. Higher origination and extinction rates in larger
mammals. *Proceedings of the National Academy of Sciences* 105:6097.
- 642 Liow, L. H. and J. D. Nichols, 2010. Estimating rates and probabilities of
origination and extinction using taxonomic occurrence data:
644 Capture-mark-recapture (CMR) approaches. *Quantitative Methods in
Paleobiology* Pp. 81–94.
- 646 Lynch, M., 1991. Methods for the analysis of comparative data in evolutionary
biology. *Evolution* 45:1065–1080.
- 648 Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings:
the kinetics of mass extinction and origination. *Science* 326:1106–9.
- 650 Munnecke, A., M. Calner, D. A. T. Harper, and T. Servais, 2010. Ordovician
and Silurian sea-water chemistry, sea level, and climate: A synopsis.
652 *Palaeogeography, Palaeoclimatology, Palaeoecology* 296:389–413.
- Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range
654 predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology*
39:360–372.
- 656 ———, 2015. Interdependence of specialization and biodiversity in Phanerozoic
marine invertebrates. *Nature communications* 6:6602.
- 658 Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental
than evolvability. *PloS one* 7:e38025.
- 660 Payne, J. L. and S. Finnegan, 2007. The effect of geographic range on

- extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences* 104:10506–11.
- 662
- Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain, 2014. Metabolic dominance of bivalves predates brachiopod diversity decline by more than 150 million years. *Proceedings of the Royal Society B* 281:20133122.
- 664
- Peters, S. E., 2008. Environmental determinants of extinction selectivity in the fossil record. *Nature* 454:626–9.
- 666
- Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law. *Paleobiology* 1:82–96.
- 668
- , 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- , 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- 670
- , 1994. The role of extinction in evolution. *Proceedings of the National Academy of Sciences* 91:6758–6763.
- 672
- Royle, J. A. and R. M. Dorazio, 2008. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Elsevier, London.
- 674
- Sepkoski Jr., J. J., 1981. A factor analytic description of the Phanerozoic marine fossil record. *Paleobiology* 7:36–53.
- 676
- Sheehan, P., 2001. The late Ordovician mass extinction. *Annual Review of Earth and Planetary Sciences* 29:331–364.
- 678
- Simpson, C., 2006. Levels of selection and large-scale morphological trends. Ph.D. thesis, University of Chicago.
- 680
- Simpson, C. and P. G. Harnik, 2009. Assessing the role of abundance in marine bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.
- 682

- Simpson, G. G., 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.
- , 1953. *The Major Features of Evolution*. Columbia University Press, New York.
- Smits, P. D., 2015. Expected time-invariant effects of biological traits on mammal species duration. *Proceedings of the National Academy of Sciences* 112:13015–13020.
- Solow, A. and W. Smith, 1997. On fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 23:271–277.
- Stan Development Team, 2014. Stan: A c++ library for probability and sampling, version 2.5.0. URL <http://mc-stan.org/>.
- Tomiya, S., 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist* 182:196–214.
- Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- , 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- Vehtari, A. and A. Gelman, 2015. Pareto Smoothed Importance Sampling. arXiv:1507.02646 .
- Vehtari, A., A. Gelman, and J. Gabry, 2015a. loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. URL <https://github.com/jgabry/loo>. R package version 0.1.3.
- , 2015b. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. arXiv:1507.04544 .
- Vilhena, D. A., E. B. Harris, C. T. Bergstrom, M. E. Maliska, P. D. Ward, C. A. Sidor, C. A. E. Strömberg, and G. P. Wilson, 2013. Bivalve network reveals

- latitudinal selectivity gradient at the end-Cretaceous mass extinction.
710 Scientific Reports 3:1790.
- Wagner, P. J. and J. D. Marcot, 2013. Modelling distributions of fossil sampling
712 rates over time, space and taxa: assessment and implications for
macroevolutionary studies. Methods in Ecology and Evolution 4:703–713.
- 714 Wang, S. C., 2003. On the continuity of background and mass extinction.
Paleobiology 29:455–467.
- 716 Wang, S. C. and C. R. Marshall, 2004. Improved confidence intervals for
estimating the position of a mass extinction boundary. Paleobiology 30:5.
- 718 Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and
Widely Applicable Information Criterion in Singular Learning Theory.
720 Journal of Machine Learning Research 11:3571–3594.