

**Title:** How macroecology affects macroevolution: the interplay  
2 between extinction intensity and trait-dependent extinction in  
brachiopods.

4 **Running title:** Trait-dependent extinction in brachiopods

**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

**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 decreases? Using a hierarchical Bayesian approach, I develop a model of how the effects of biological traits on extinction risk can vary with respect to extinction intensity, origination cohort (i.e. 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 environmental generalized versus specialized on taxonomic survival by allowing environmental preference to have a nonlinear effect on duration. My analytical framework eschews the traditional distinction between background and mass extinction, and instead considers extinction intensity as a continuum. I find that the cohort-specific effects of geographic range and environmental preference are negatively correlated with baseline extinction intensity. I also infer that as extinction intensity increases, while both effects will increase the change in magnitude of the effect environmental preference will be greater than the change in the effect of geographic range. Additionally, I find support for greater survival of environmental generalists versus specialists in all origination cohorts, though taxa favoring epicontinental environments have a slightly greater duration than open-ocean favoring taxa. These results support the conclusion that for Paleozoic brachiopods, as extinction intensity increases overall extinction selectivity increases.

<sup>38</sup> **1 Introduction**

How do biological traits affect extinction risk? Biological traits the set of  
<sup>40</sup> adaptations a taxon has to the set of all biotic–biotic and biotic–abiotic  
interactions that a taxon can experience, which is their adaptive zone (Simpson,  
<sup>42</sup> 1944). Jablonski (1986) observed that for bivalves at the end Cretaceous mass  
extinction event, the effects of some biological traits on taxonomic survival  
<sup>44</sup> decreased. However, this pattern was not the case for the effect of geographic  
range on survival (Jablonski, 1986, Payne and Finnegan, 2007). There are  
<sup>46</sup> multiple possible macroevolutionary mechanisms which may underlie this  
pattern: the effect of geographic range on survival remains constant and those of  
<sup>48</sup> other biological traits decrease, the effect of geographic range on survival  
increases and those of other biological traits stay constant, or the effects of all  
traits decrease potentially by different degrees.  
<sup>50</sup>

While Jablonski (1986) phrased his conclusions in terms of background versus  
<sup>52</sup> mass extinction, these states are not distinguishable in terms of extinction rate  
alone; my analysis treats the time period analyzed as part of the same  
<sup>54</sup> continuum (Payne and Finnegan, 2007, Simpson and Harnik, 2009, Wang, 2003).  
Additionally, in order to test the proposed macroevolutionary mechanism behind  
<sup>56</sup> the Jablonski (1986) scenario; not only do the taxon trait effects needs to be  
modeled, but the correlation between trait effects need to be modeled as well.  
<sup>58</sup> Conceptually, taxon survival can be considered an aspect of “taxon fitness”  
along with expected lineage specific branching/origination rate (Cooper, 1984,  
<sup>60</sup> 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  
<sup>62</sup> 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, 2010).

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 modeling approach (Gelman et al., 2013, Gelman and Hill, 2007, Klein and Moeschberger, 2003), 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 biological trait for estimating differences in extinction risk at nearly all times, 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).

Epicontinental seas are a shallow-marine environment where the ocean has spread over the craton of a continental shelf with a depth typically less than 100m. In contrast, open-ocean coastline environments have much greater

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

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

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

<sup>116</sup> Body size, measured as shell length, is also considered as a trait that may potentially influence extinction risk (Harnik, 2011, Payne et al., 2014). Body  
<sup>118</sup> size is a proxy for metabolic activity and other correlated life history traits (Payne et al., 2014). Harnik et al. (2014) analyzed the effect of body size  
<sup>120</sup> selectivity in Devonian brachiopods in both a phylogenetic and non-phylogenetic context; finding that body size was not found to be associated with differences in taxonomic duration. It has also been found that, at least in the case of some bivalve 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 body size has any effect on brachiopod taxonomic survival it is very small.  
<sup>126</sup>

It is well known that, given the incompleteness of the fossil record, the observed 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, 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 being sampled and not any of the effects of the covariates of interest. In this case, for sampling to be a confounding factor there must be consistent relationship between the quality of sampling of a taxon and its apparent duration (e.g. greater sampling, longer duration). If there is no relationship between sampling and duration then interpretation can be made clearly; while observed durations are obviously truncated true durations, a lack of a relationship would indicate that the amount and form of this truncation is not a major determinant of the taxon's apparent duration.

<sup>140</sup> **2 Materials and Methods**

**2.1 Fossil occurrence information**

<sup>142</sup> The brachiopod dataset analyzed here was sourced from the Paleobiology Database (<http://www.paleodb.org>) which was then filtered based on  
<sup>144</sup> taxonomic (Rhynchonelliformea: Rhynchonellata, Chileata, Obolellida,  
Kutorginida, Strophomenida, Spiriferida)), temporal (post-Cambrian Paleozoic),  
<sup>146</sup> stratigraphic, and other occurrence information used in this analysis. Analyzed occurrences were restricted to those with paleolatitude and paleolongitude  
<sup>148</sup> coordinates, assignment to either epicontinental or open-ocean environment, and belonging to a genus present in the body size dataset (Payne et al., 2014).  
<sup>150</sup> Epicontinental versus open-ocean assignments for each fossil occurrence are partially based on those from Miller and Foote (2009), with additional  
<sup>152</sup> occurrences assigned similarly (Miller and Foote, personal communication). These filtering criteria are very similar to those from Foote and Miller (2013)  
<sup>154</sup> with an additional constraint of being present in the body size data set from Payne et al. (2014).  
<sup>156</sup> Fossil occurrences were analyzed at the genus level which is common for paleobiological, macroevolutionary and macroecological studies of marine  
<sup>158</sup> invertebrates (Alroy, 2010, Foote and Miller, 2013, Harnik et al., 2012, Kiessling and Aberhan, 2007, Miller and Foote, 2009, Nürnberg and Aberhan, 2013, 2015,  
<sup>160</sup> Payne and Finnegan, 2007, Simpson and Harnik, 2009, Vilhena et al., 2013). While species diversity dynamics are frequently of much greater interest than  
<sup>162</sup> those of higher taxa (though see Foote 2014, Hoehn et al. 2015), the nature of the fossil record makes accurate and precise taxonomic assignments at the  
<sup>164</sup> species level difficult for all occurrences. As such, the choice to analyze genera as opposed to species was in order to assure a minimum level of confidence and

<sup>166</sup> accuracy in the data analyzed here.

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

Geographic range was calculated using an occupancy approach corrected for  
<sup>180</sup> incomplete sampling. First, all occurrences were projected onto an equal-area  
cylindrical map projection. Each occurrence was then assigned to one of the  
<sup>182</sup> cells from a  $70 \times 34$  regular raster grid placed on the map. Each grid cell  
represents approximately 250,000 km<sup>2</sup>. The map projection and regular lattice  
<sup>184</sup> were made using shape files from <http://www.naturalearthdata.com/> and the  
`raster` package for R (Hijmans, 2015). For each stage, the total number of  
<sup>186</sup> occupied grid cells was calculated. Then, for each temporal bin, the relative  
occurrence probability of the observed taxa was calculated using the JADE  
<sup>188</sup> method developed by Chao et al. (2015). This method accounts for the fact that  
taxa with an occupancy of 0 cannot be observed which means that occupancy is  
<sup>190</sup> estimated from a truncated Binomial distribution. This correction is critical  
when comparing occupancies from different times with different geographic  
<sup>192</sup> sampling. Finally, for each genus, the mean relative occurrence probability was

calculated as the average of that genus' occurrence probabilities for all stages it  
194 was sampled.

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

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

200 where  $v$  is the percent of the distribution defined in equation 1 less than or  
equal to  $\theta_i$ . The Beta distribution is used here because it is a continuous  
202 distribution bounded at 0 and 1, which is idea for modeling percentages.

Body size, measured as shell length, was sourced directly from Payne et al.  
204 (2014). These measurements were made from brachiopod taxa figured in the  
*Treatise on Invertebrate Paleontology* (Williams et al., 2007).

206 The sampling record for individual taxa was calculated using the standard gap  
statistic (Foote, 2000a, Foote and Raup, 1996). The gap statistic is calculated as  
208 the number of stages in which the taxon was sampled minus two divided by the  
duration of the taxon minus two. Subtracting two from both the numerator and  
210 denominator is because the first and last appearance stages are by definition  
sampled. The minimum duration for which a gap statistic can be calculated is  
212 three stages.

Prior to analysis, geographic range was logit transformed and body size was  
214 natural-log transformed; both of these transformations make these variables  
defined for the entire real line. All covariates were then standardized by  
216 subtracting the mean from all values and dividing by twice its standard

deviation, which follows Gelman and Hill (2007). This standardization means  
218 that all regression coefficients are comparable as the expected change per 1-unit  
change in the rescaled covariates.  $D$  is defined as the total number of covariate  
220 effects included in the model plus one for the intercept term.

## 2.2 Analytical approach

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

Genus durations were assumed to follow a Weibull distribution which allows for  
238 age-dependent extinction (Klein and Moeschberger, 2003):  $y \sim \text{Weibull}(\alpha, \sigma)$ .  
The Weibull distribution has two parameters: scale  $\sigma$ , and shape  $\alpha$ . When  
240  $\alpha = 1$ ,  $\sigma$  is equal to the expected duration of any taxon.  $\alpha$  is a measure of the  
effect of age on extinction risk where values greater than 1 indicate that

<sup>242</sup> extinction risk increases with age, and values less than 1 indicate that  
extinction risk decreases with age. Note that the Weibull distribution is  
<sup>244</sup> equivalent to the exponential distribution when  $\alpha = 1$ .

The scale parameter  $\sigma$  was modeled as a regression following Kleinbaum and  
<sup>246</sup> 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  
<sup>248</sup> observation,  $X$  is a  $N \times D$  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  
<sup>250</sup> coefficients.

Each of the rows of matrix  $B$  are modeled as realizations from a multivariate  
<sup>252</sup> normal distribution with length  $D$  location vector  $\mu$  and  $J \times J$  covariance  
matrix  $\Sigma$ :  $B_j \sim \text{MVN}(\mu, \Sigma)$ . The covariance matrix was then decomposed into  
<sup>254</sup> a length  $J$  vector of scales  $\tau$  and a  $J \times J$  correlation matrix  $\Omega$ , defined

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

where “diag” indicates a diagonal matrix.

<sup>256</sup> The elements of  $\mu$  were given independent normally distributed priors. The  
effects of geographic range size and the breadth of environmental preference  
<sup>258</sup> were given informative priors reflecting the previous findings while the others  
were given weakly informative favoring no effect. The correlation matrix  $\Omega$  was  
<sup>260</sup> given an almost uniform LKJ distributed prior following (Team, 2016). These

priors are defined

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

- <sup>262</sup> The shape parameter  $\alpha$  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  
<sup>264</sup> shared as normal distribution such that  $a_j \sim \mathcal{N}(0, \sigma_a)$ .  $\alpha$ ,  $\sigma_\alpha$  and  $\sigma_a$  were given  
weakly informative normal and half-Cauchy priors as appropriate:

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

- <sup>266</sup> The above model is for all taxa without sampling as a covariate. In order to  
determine if sampling is acting as a confounding factor in this analysis, an  
<sup>268</sup> additional model was developed because sampling was only estimated for taxa  
with a duration of three or more which creates a left-truncated distribution of  
<sup>270</sup> durations (Klein and Moeschberger, 2003). The sampling statement and  
log-probability for a left-truncated Weibull distribution, truncated at time  $Y$

<sup>272</sup> (e.g. three), is

$$\begin{aligned} p(y|\theta) &= \frac{\text{Weibull}(y, \alpha, \sigma)}{1 - \text{Weibull}_{cdf}(Y, \alpha, \sigma)} \\ p(y|\theta) &= \frac{\text{Weibull}(y, \alpha, \sigma)}{\text{Weibull}_{ccdf}(Y, \alpha, \sigma)} \end{aligned} \quad (6)$$

$$\log(p(y|\theta)) = \log(\text{Weibull}(y, \alpha, \sigma)) - \log(\text{Weibull}_{ccdf}(Y, \alpha, \sigma)).$$

Note that cdf stands for cumulative density function and ccdf is the  
<sup>274</sup> complementary cumulative density function.

The definition of  $\sigma$  (Eq. 2) is then further modified so that  $X$ , the matrix of  
<sup>276</sup> covariates, and  $B$ , the matrix of regression coefficients, now include an  
additional column for the sampling estimates and the cohort-specific effects of  
<sup>278</sup> sampling. This addition then modifies the dimensions of  $\mu$  and  $\Sigma$ ; the new  
group-level effect of  $\mu_s$  is given a weakly informative prior:  $\mu_s \sim \mathcal{N}(0, 1)$ .

<sup>280</sup> For this left-truncated model, I've excluded one observation that is  
right-censored with a duration equal to the truncation point; the second line of  
<sup>282</sup> equation 6 becomes  $p(y|\theta) = \text{Weibull}_{ccdf}(y, \alpha, \sigma)/\text{Weibull}_{ccdf}(Y, \alpha, \sigma)$  which  
yields a log-probability of 0 which is undefined.

<sup>284</sup> 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  
<sup>286</sup> Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic  
programming language Stan (Stan Development Team, 2014). The posterior  
<sup>288</sup> 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  
<sup>290</sup> 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  
<sup>292</sup> convergence. Convergence means that the chains are approximately stationary  
and the samples are well mixed (Gelman et al., 2013).

<sup>294</sup> Model adequacy was evaluated using a couple of posterior predictive checks.  
<sup>295</sup> Posterior predictive checks are a means for understanding model fit or adequacy  
<sup>296</sup> where the basic idea is that replicated data sets simulated from the fitted model  
<sup>297</sup> should be similar to the original data and systematic differences between the  
<sup>298</sup> simulations and observations indicate weaknesses of the model fit (Gelman  
<sup>299</sup> et al., 2013). For both approaches used here, 1000 posterior predictive datasets  
<sup>300</sup> were generated from 1000 unique draws from the posterior distribution of each  
<sup>301</sup> parameter. The two posterior predictive checks used in this analysis are a  
<sup>302</sup> comparison of a non-parameteric estimate of the survival function  $S(t)$  from the  
<sup>303</sup> empirical dataset to the non-parameteric estimates of  $S(t)$  from the 1000  
<sup>304</sup> posterior predictive datasets, and comparison of the observed genus durations to  
<sup>305</sup> the average posterior predictive estimate of  $\log(\sigma)$  (Eq. 2). The former is to see  
<sup>306</sup> 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.

### <sup>308</sup> 3 Results

The model with greatest estimated out-of-sample predictive accuracy, as  
<sup>310</sup> estimated via both WAIC and LOOIC, has both constant  $\alpha$  and no sampling  
<sup>311</sup> covariate (Table 1). The effects of either of these potential model aspects are  
<sup>312</sup> inferred to be too small to account for the additional model complexity  
necessary to estimate them; because of this, only the results from the  
<sup>314</sup> WAIC/LOOIC best model will be presented here.

Comparison of the posterior predictive estimates of  $S(t)$  to the empirical  
<sup>316</sup> 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  
<sup>318</sup> additional posterior predictive comparison where most estimates are not

systematically biased except for a consistent under-estimate of  $\log(\sigma)$  for older  
320 taxa (Fig. 2). The results of both posterior predictive checks indicate that, for  
the majority of observations, model fit is generally not biased.

322 The cohort-level estimate of the effect of geographic range size indicates that as  
a taxon's geographic range increases, that taxon's duration is expected to  
324 increase (Table 2). Given the estimates of  $\mu_r$  and  $\tau_r$ , there is a less than 0.008%

( $\pm 0.05$  SD) probability that this relationships would be reversed

326 ( $\Pr(\mathcal{N}(\mu_r, \tau_r) > 0)$ ). The between-cohort variance  $\tau_r$  is the lowest of all the  
regression coefficients (Table 2).

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

332 The group-level estimate of the effect of environmental preference is estimated  
from both  $\mu_v$  and  $\mu_{v^2}$ .

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

336 given the estimate of between-cohort variance  $\tau_v$ , there is approximately 19%  
( $\pm 8$  SD) probability that, for any given cohort, taxa favoring open-ocean

338 environments would have a greater expected duration than taxa favoring  
epicontinental environments ( $\Pr(\mathcal{N}(\mu_v, \tau_v) > 0)$ ).

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

The cohort-specific estimates of all the regression coefficients demonstrate a lot  
344 of between cohort variance, with no obvious trends. As indicated in Table 2 and

detectable visually (Fig. 4), the between-cohort estimates for  $\beta_0$ ,  $\beta_r$ , and  $\beta_m$  all  
346 have much lower variance than the between-cohort estimates of both  $\beta_v$  and  $\beta_{v^2}$ .

While most cohort-specific estimates are very similar to the overall cohort-level  
348 estimate, there are a few notable excursions away from the overall mean (Fig.  
4). There are simultaneous excursions in both  $\beta_0$  and  $\beta_v$  for cohorts originating  
350 in the Givetian (387-382 My) and Frasnian (382-372 My) stages; both of which  
directly precede the late Devonian mass extinction event at the  
352 Frasnian/Famennian boundary. These cohorts are marked by both a high  
extinction intensity and an increase in expected duration for taxa favoring  
354 epicontinental environments over open-ocean ones; this is consistent with the  
results of Miller and Foote (2009).

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

The cohort-specific relationships between environmental preference and  $\log(\sigma)$   
366 were calculated from the estimates of  $\beta_0$ ,  $\beta_v$ , and  $\beta_{v^2}$  (Fig. 5) and reflect how  
these three parameters act in concert and not just individually (Fig. 4). Beyond  
368 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  
370 like the overall expected relationship and has the most sharply curved  
appearance due to a high estimate  $\beta_{v^2}$  (Fig. 4). This cohort has the biggest

372 difference in extinction risk between environmental generalists and specialists.

373 The cohorts originating during the Emsian (407-393 My) and Frasnian (382 -

374 372 My) are tied for second in 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, most of these cohorts originate during the Silurian

378 through the Early Devonian range identified earlier as having lower expected

379 extinction intensity than what is expected from the group-level estimate.

380 The correlations of the cohort-specific estimates of the regression coefficients are

381 estimated as the off-diagonal elements of the correlation matrix  $\Omega$ . Only two of

382 the elements of  $\Omega$  are distinguishable from 0: the correlation of  $\beta_0$  (extinction

383 intensity) with both  $\beta_r$  and  $\beta_v$  (Fig. 6).

384 There is an approximate 86% probability that the cohort-specific estimates of

385 baseline extinction intensity  $\beta_0$  and the effect of geographic range  $\beta_r$  are

386 negatively correlated; this means that for cohorts experiencing a lower

387 extinction intensity ( $\beta_0$  decreases), the magnitude of the effect of geographic

388 range is expected to decrease as well, and *vice versa*.

389 Similarly, there is an approximate 99.9% probability that the cohort-specific

390 estimates of  $\beta_0$  and  $\beta_v$  are negatively correlated; this means that as extinction

391 intensity increases it is expected that epicontinental taxa become more favored

392 over open-ocean environments (i.e. as  $\beta_0$  increases,  $\beta_v$  decreases). Additionally,

393 there is only a 19% ( $\pm 8.1$  SD) probability that, for any given cohort,

394 open-ocean environments will be favored ( $\Pr(|\mu_v| - 2\tau_v > 0)$ ).

395 There is only an approximate 68% probability that  $\beta_r$  and  $\beta_v$  are positively

396 correlated. This lack of cross-correlation may be due in part to the much higher

397 between-cohort variance of the effect of environmental preference  $\tau_v$  than the

<sup>398</sup> very small between-cohort variance in the effect of geographic range  $\tau_r$  (Table  
2); the effect of geographic range might simply not vary enough relative to the  
<sup>400</sup> much noisier environmental preference.

## 4 Discussion

<sup>402</sup> The generating observation behind this study was that for bivalves at the end  
Cretaceous mass extinction event, the only biological trait that was found the  
<sup>404</sup> affect extinction risk was geographic range while traits that had previously been  
beneficial had no effect (Jablonski, 1986). This observation raises two linked  
<sup>406</sup> questions: how does the effect of geographic range change with changing  
extinction intensity, and how does the effect of other biological traits change  
<sup>408</sup> with changing extinction intensity?

I find that as intensity increases ( $\beta_0$  decreases), the magnitude of the effect of  
<sup>410</sup> geographic range increase. I also find that as intensity increases, the effect of  
favoring epicontinental environments of open-ocean environments is expected to  
<sup>412</sup> be increase; this is consistent with the results of Miller and Foote (2009). There  
is no evidence for a correlation between the effect of geographic range and  
<sup>414</sup> environmental preference. However, the between-cohort variance in effect of  
geographic range is much less then the between-cohort variance of the effect of  
<sup>416</sup> environmental preference. As stated earlier, this may underlie the lack of  
correlation between these two effects.

<sup>418</sup> Additionally, the lower between-cohort variance in the effect of geographic range  
versus that higher between-cohort variance implies that for cohorts with a  
<sup>420</sup> greater than average extinction intensity, the difference in the effect geographic  
range and the group-level effect of geographic range is expected to be smaller  
<sup>422</sup> than the difference between the effect of environmental preference and the

group-level effect of environmental preference.

424 I find consistent support for the “survival of the unspecialized,” with respect to  
epicontinental versus open-ocean environmental preference, as a time-invariant  
426 generalization of brachiopod survival; taxa with intermediate environmental  
preferences are expected to have lower extinction risk than taxa specializing in  
428 either epicontinental or open-ocean environments (Fig. 3), though the curvature  
of the relationship varies from rather shallow to very peaked (Fig. 5). However,  
430 this relationship is not symmetric about 0, as taxa favoring epicontinental  
environments are expected to have a greater duration than taxa favoring  
432 open-ocean environments. This description of environment only describes one  
major aspect of a taxon’s environmental context, with factors such as  
434 bathymetry and temperature being further descriptors of a taxon’s adaptive  
zone (Harnik, 2011, Harnik et al., 2012, Heim and Peters, 2011, Nürnberg and  
436 Aberhan, 2013); inclusion of these factors in future analyses would potentially  
improve our understanding of the “survival of the unspecialized” hypothesis  
438 (Simpson, 1944).

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

<sup>450</sup> duration in relation to environmental availability.

In this study I used the average quality of the record for the duration of a taxon  
<sup>452</sup> as a measure of the effect of sampling with the goal of determining if sampling  
is acting as a confounding variable where taxa with shorter or longer durations  
<sup>454</sup> are associated with differences in sampling regime. This covariate, however, is  
not present in the WAIC or LOOIC best model (Table 1). I conclude then that  
<sup>456</sup> the effect of sampling on duration is at most small and is not acting as a  
confounding factor.

<sup>458</sup> An alternative approach for specifically modeling survival that can take into  
account imperfect observation than the method used here is the  
<sup>460</sup> Cormack-Jolly-Seber (CJS) model (Liow et al., 2008, Liow and Nichols, 2010,  
Royle and Dorazio, 2008, Tomiya, 2013). This model is a type of hidden Markov  
<sup>462</sup> 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  
<sup>464</sup> 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  
<sup>466</sup> 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  
<sup>468</sup> time  $t + 1$ , I model the distribution of taxon durations (lifetime) which then  
decomposes into a function describing survival ( $S(t)$ ) and a function describing  
<sup>470</sup> instantaneous extinction rate with taxon age (Klein and Moeschberger, 2003,  
Royle and Dorazio, 2008).

<sup>472</sup> The use of genera as the unit of the study and how to exactly interpret the  
effects of the biological traits is an important question. For example, if any of  
<sup>474</sup> the traits analyzed here are associated with increases in speciation rates, this  
might increase the duration of genera through self-renewal (Raup, 1991, 1994),  
<sup>476</sup> 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  
478 appearing to decrease generic level extinction risk by that trait increasing  
species level origination rate instead of decreasing species level extinction risk.

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

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

496 Traits like environmental preference or geographic range (Hunt et al., 2005,  
Jablonski, 1987) are most likely heritable. 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  
505 additional individual-level hierarchical effect independent of origination cohort  
506 would allow for determining how much of the observed variability is due to  
507 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  
511 traditional separation between background and mass extinction. I find that  
512 cohort extinction intensity is negatively correlated with both the cohort-specific  
513 effects of geographic range and environmental preference. These results imply  
514 that as extinction intensity increases ( $\beta_0$ ) increases, it is expected that both  
515 effects will increase in magnitude. However, the change in effect of  
516 environmental preference is expected to be greater than the change in the effect  
517 of geographic range. Additionally, I find support for greater survival in  
518 environmental generalists over specialists in all origination cohorts analyzed;  
519 this is consistent with the long standing “survival of the unspecialized”  
520 hypothesis (Baumiller, 1993, Liow, 2004, 2007, Nürnberg and Aberhan, 2013,  
521 2015, Simpson, 1944, 1953, Smits, 2015). I also find that taxa favoring  
522 epicontinental environments are expected to have a slightly greater duration  
523 than open-ocean favoring taxa. The results of this analysis support the  
524 conclusion that for Paleozoic brachiopods, as extinction intensity increases  
525 overall extinction selectivity is expected to increase as well.

## 526 Acknowledgements

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, R. Ree, and G. Slater  
527 for helpful discussion during the conception of this study. I'd also like to thank

D. Bapst, N. Pierrehumbert and M. Villarosa Garcia for additional comments.  
530 An additional thank you to A. Miller for the epicontinental versus open-ocean  
assignments. Finally, thank you to the reviewers for their helpful comments that  
532 improved this manuscript. This entire study would have not been possible  
without the 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, F. Fürsich, N. Heim, A. Hendy, S. Holland, L. Ivany, W. Kiessling,  
536 B. Kröger, A. McGowan, T. Olszewski, P. Novack-Gottshall, M. Patzkowsky, M.  
Uhen, L. Villier, and P. Wager. This work was supported by a NASA  
538 Exobiology grant (NNX10AQ446) to A. Miller and M. Foote. I declare no  
conflicts of interest. This is Paleobiology Database publication XXX.

## 540 References

- Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation  
542 to the mode of stabilization on the substrate. *Palaeogeography, Palaeoclimatology, Palaeoecology* 21:209–226.
- 544 Alroy, J., 2010. The Shifting Balance of Diversity Among Major Marine Animal Groups. *Science* 329:1191–1194.
- 546 ———, 2014. A simple Bayesian method of inferring extinction. *Paleobiology* 40:584–607.
- 548 Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. *Paleobiology* 19:304–321.
- 550 Chao, A., T. C. Hsieh, R. L. Chazdon, R. K. Colwell, N. J. Gotelli, and B. D. Inouye, 2015. Unveiling the species-rank abundance distribution by  
552 generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- Cooper, W. S., 1984. Expected time to extinction and the concept of  
554 fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote, 2016.  
556 Greenhouseicehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *Proceedings of the National  
558 Academy of Sciences P.* 201519092.
- Edwards, D. and U. Fanning, 1985. Evolution and environment in the late  
560 Silurian–early Devonian: the rise of pteridophytes. *Philosophical Transactions 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*  
564 8:139–42.

- Felsenstein, J., 1985. Phylogenies and the comparative method. American  
566 Naturalist 125:1–15.
- Finnegan, S., N. A. Heim, S. E. Peters, and W. W. Fischer, 2012. Climate  
568 change and the selective signature of the Late Ordovician mass extinction.  
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:  
general problems. Paleobiology 26:74–102.
- 572 ———, 2000b. Origination and extinction components of taxonomic diversity:  
Paleozoic and post-Paleozoic dynamics. Paleobiology 26:578–605.
- , 2014. Environmental controls on geographic range size in marine animal  
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.
- Gelman, A. and J. Hill, 2007. Data Analysis using Regression and  
586 Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- Harnik, P. G., 2011. Direct and indirect effects of biological factors on

- 588 extinction risk in fossil bivalves. *Proceedings of the National Academy of Sciences of the United States of America* 108:13594–9.
- 590 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.
- 594 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.
- 596 Heim, N. A. and S. E. Peters, 2011. Regional environmental breadth predicts geographic range and longevity in fossil marine genera. *PloS one* 6:e18946.
- 598 Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL <http://CRAN.R-project.org/package=raster>. R package version 2.3-24.
- 600 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* .
- 604 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.
- 606 Hopkins, M. J., C. Simpson, and W. Kiessling, 2014. Differential niche dynamics among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- 608 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 extinctions: the alternation of macroevolutionary regimes. *Science* 231:129–133.

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

- 636 range and morphological variability contribute to taxon longevity? Global  
Ecology and Biogeography 16:117–128.
- 638 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  
640 mammals. Proceedings of the National Academy of Sciences 105:6097.
- Liow, L. H. and J. D. Nichols, 2010. Estimating rates and probabilities of  
642 origination and extinction using taxonomic occurrence data:  
Capture-mark-recapture (CMR) approaches. Quantitative Methods in  
644 Paleobiology Pp. 81–94.
- Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings:  
646 the kinetics of mass extinction and origination. Science 326:1106–9.
- Munnecke, A., M. Calner, D. A. T. Harper, and T. Servais, 2010. Ordovician  
648 and Silurian sea-water chemistry, sea level, and climate: A synopsis.  
Palaeogeography, Palaeoclimatology, Palaeoecology 296:389–413.
- 650 Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range  
predict diversity dynamics in marine Mesozoic bivalves. Paleobiology  
652 39:360–372.
- , 2015. Interdependence of specialization and biodiversity in Phanerozoic  
654 marine invertebrates. Nature communications 6:6602.
- Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental  
656 than evolvability. PloS one 7:e38025.
- Payne, J. L. and S. Finnegan, 2007. The effect of geographic range on  
658 extinction risk during background and mass extinction. Proceedings of the  
National Academy of Sciences 104:10506–11.
- 660 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  
662 million years. *Proceedings of the Royal Society B* 281:20133122.
- Peters, S. E., 2008. Environmental determinants of extinction selectivity in the  
664 fossil record. *Nature* 454:626–9.
- Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law.  
666 *Paleobiology* 1:82–96.
- \_\_\_\_\_, 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- 668 \_\_\_\_\_, 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- \_\_\_\_\_, 1994. The role of extinction in evolution. *Proceedings of the National  
670 Academy of Sciences* 91:6758–6763.
- Royle, J. A. and R. M. Dorazio, 2008. Hierarchical modeling and inference in  
672 ecology: the analysis of data from populations, metapopulations and  
communities. Elsevier, London.
- 674 Sepkoski, 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.

- 684 ———, 1953. *The Major Features of Evolution*. Columbia University Press,  
New York.
- 686 Smits, P. D., 2015. Expected time-invariant effects of biological traits on  
mammal species duration. *Proceedings of the National Academy of Sciences*  
688 112:13015–13020.
- 690 Solow, A. and W. Smith, 1997. On fossil preservation and the stratigraphic  
ranges of taxa. *Paleobiology* 23:271–277.
- 692 Stan Development Team, 2014. Stan: A c++ library for probability and  
sampling, version 2.5.0. URL <http://mc-stan.org/>.
- Team, S. D., 2016. Stan Modeling Language Users Guide and Reference Manual.
- 694 Tomiya, S., 2013. Body Size and Extinction Risk in Terrestrial Mammals Above  
the Species Level. *The American Naturalist* 182:196–214.
- 696 Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.  
———, 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- 698 Vehtari, A. and A. Gelman, 2015. Pareto Smoothed Importance Sampling.  
arXiv:1507.02646 .
- 700 Vehtari, A., A. Gelman, and J. Gabry, 2015a. loo: Efficient leave-one-out  
cross-validation and WAIC for Bayesian models. URL  
702 <https://github.com/jgabry/loo>. R package version 0.1.3.
- , 2015b. Practical Bayesian model evaluation using leave-one-out  
704 cross-validation and WAIC. arXiv:1507.04544 .
- Vilhena, D. A., E. B. Harris, C. T. Bergstrom, M. E. Maliska, P. D. Ward, C. A.  
706 Sidor, C. A. E. Strömberg, and G. P. Wilson, 2013. Bivalve network reveals

- latitudinal selectivity gradient at the end-Cretaceous mass extinction.  
708      Scientific Reports 3:1790.
- Wagner, P. J. and J. D. Marcot, 2013. Modelling distributions of fossil sampling  
710      rates over time, space and taxa: assessment and implications for  
macroevolutionary studies. Methods in Ecology and Evolution 4:703–713.
- 712      Wang, S. C., 2003. On the continuity of background and mass extinction.  
Paleobiology 29:455–467.
- 714      Wang, S. C. and C. R. Marshall, 2004. Improved confidence intervals for  
estimating the position of a mass extinction boundary. Paleobiology 30:5.
- 716      Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and  
Widely Applicable Information Criterion in Singular Learning Theory.  
718      Journal of Machine Learning Research 11:3571–3594.
- Williams, C., C. H. C. Brunton, and S. J. Carlson, 2007. Treatise on  
720      invertebrate paleontology. Part H, Brachiopoda. Geological Society of  
America, Boulder, Colorado.

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 $\alpha$	9080.39	4823.13
constant $\alpha$ , no sampling	<b>4676.84</b>	<b>4621.95</b>
no sampling	8250.22	5563.97
full model	759172.46	109417.64

Table 2: Group-level estimates of the effects of biological traits on brachiopod generic survival.  $\mu$  values are the location parameters of the effects, while  $\tau$  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	estimate	mean	standard deviation	10%	50%	90%
$\mu_i$	intercept	-2.32	0.14	-2.50	-2.32	-2.15
$\mu_r$	geographic range	-0.76	0.11	-0.91	-0.76	-0.62
$\mu_v$	environmental preference	-0.66	0.17	-0.88	-0.66	-0.43
$\mu_{v^2}$	environmental preference	2.88	0.31	2.48	2.88	3.27
$\mu_m$	body size	0.04	0.12	-0.12	0.04	0.19
$\tau_i$	Var(intercept)	0.50	0.10	0.37	0.49	0.63
$\tau_r$	Var(geographic range)	0.27	0.13	0.11	0.26	0.45
$\tau_v$	Var(environmental preference)	0.76	0.16	0.56	0.74	0.97
$\tau_{v^2}$	Var(environmental preference)	1.24	0.33	0.84	1.21	1.67
$\tau_m$	Var(body size)	0.47	0.12	0.33	0.47	0.63

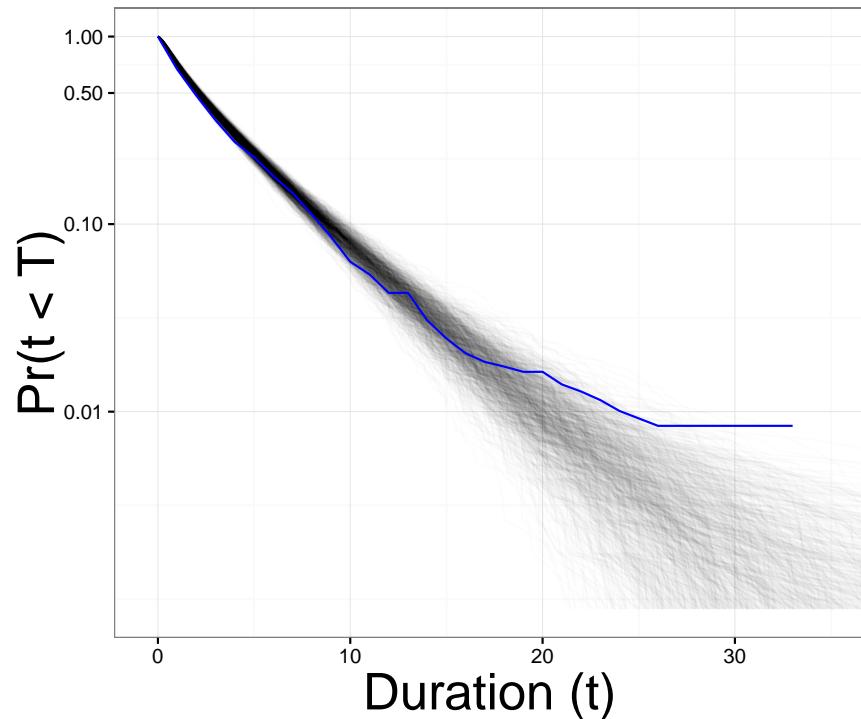


Figure 1: Comparison of the empirical estimate of  $S(t)$  (highlighted) versus estimates from 1000 posterior predictive data sets (black).  $S(t)$  corresponds to the probability that the age of a genus  $t$  is less than the genus' ultimate duration  $T$ . The vertical axis is log10 transformed.

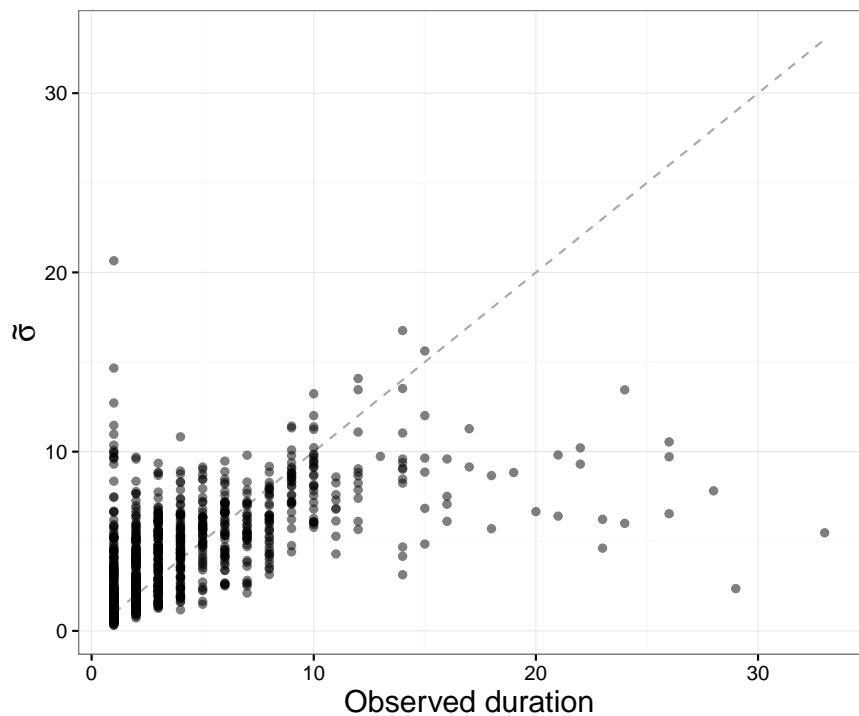


Figure 2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of  $\log(\sigma)$ . The dashed, diagonal line corresponds to  $x = y$ .

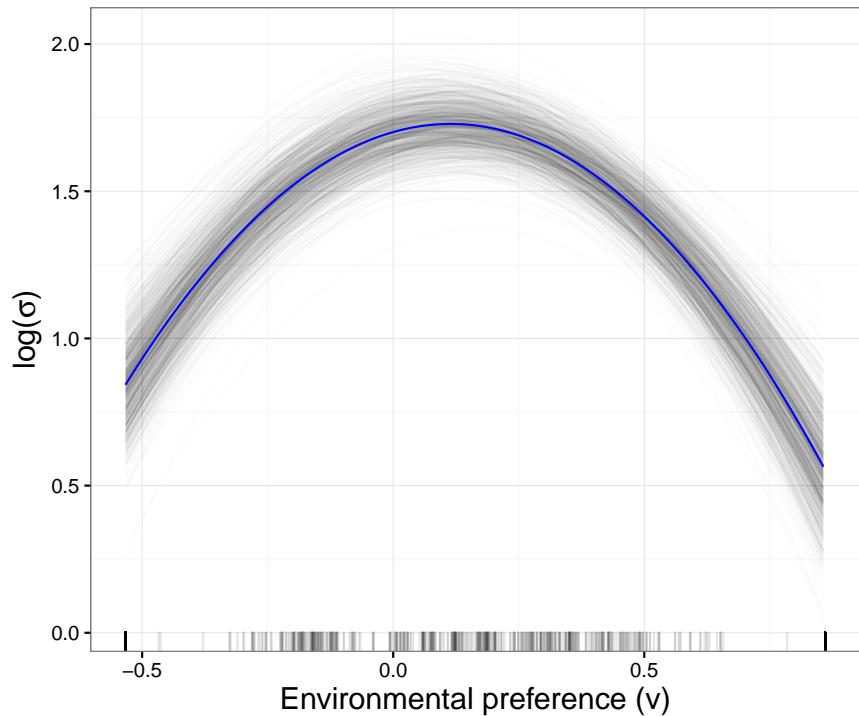


Figure 3: The overall expected relationship between environmental affinity  $v_i$  and a  $\log(\sigma)$  when  $r = 0$  and  $m = 0$ . Each grey line corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive 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.

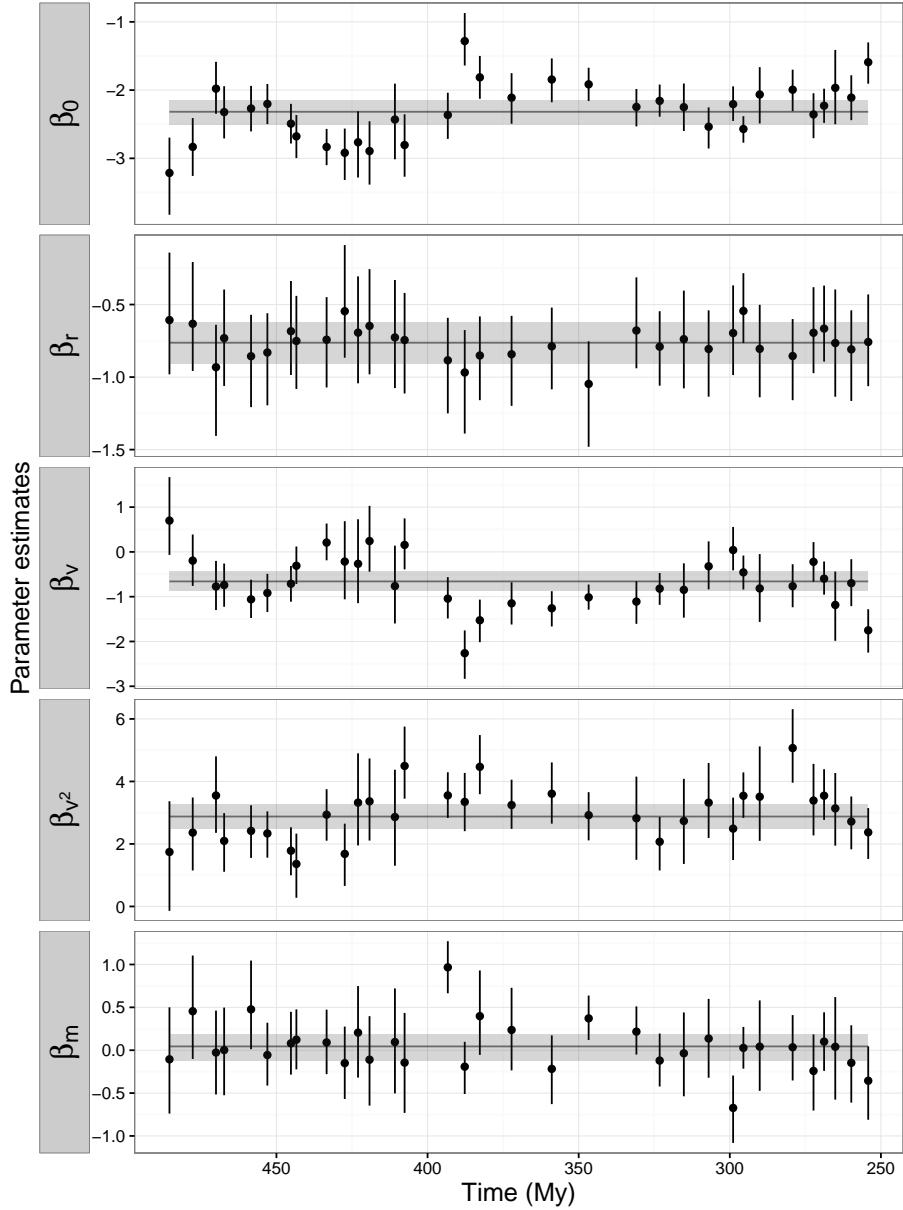


Figure 4: Comparison of cohort-specific estimates of  $\beta_0$ , the effect of geographic range on extinction risk  $\beta_r$ , the effect of environmental preference  $\beta_v$  and  $\beta_{v^2}$ , and body size  $\beta_m$ . Points correspond to the median of the cohort-specific estimate, along with 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).

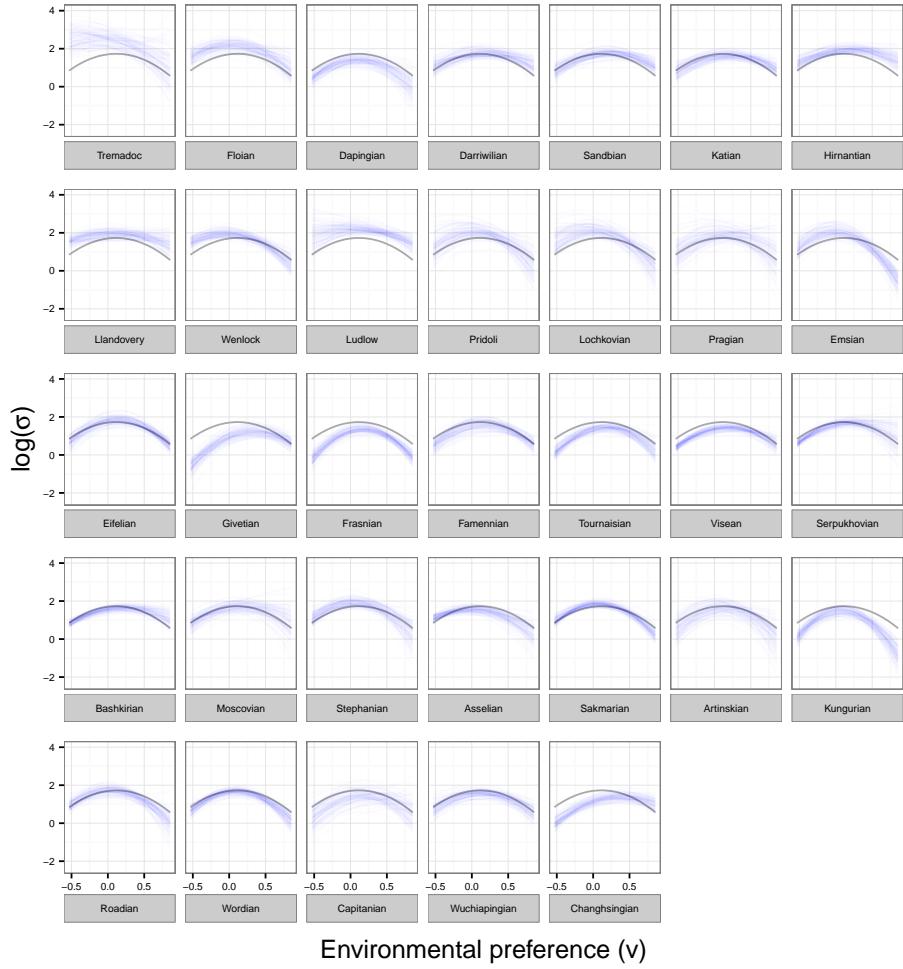


Figure 5: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on  $\log(\sigma)$  to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the 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.

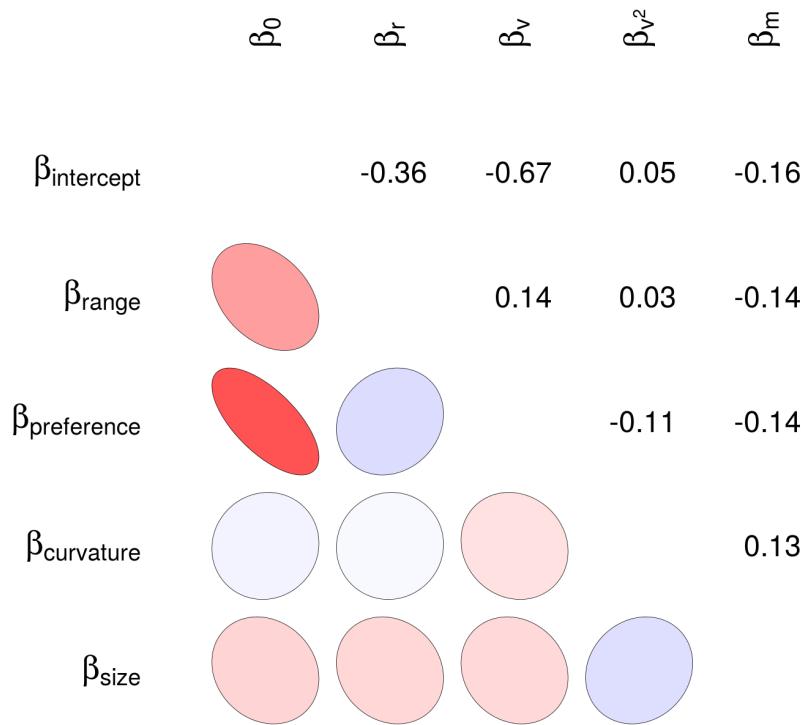


Figure 6: Mixed graphical and numerical representation of the correlation matrix  $\Omega$  of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates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