

**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 extinction rate associated with 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.

The emergent traits I analyze in relation to their patterns of Paleozoic brachiopod genus durations are geographic range, affinity for epicontinental seas versus open ocean environments, and body size. Additionally, I estimate the effects of environmental generalization versus specialization 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, though both effects will increase, the variance in magnitude of the effect of environmental preference will be greater than the variance 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? Extinction is one half of the  
<sup>40</sup> diversification process (Raup, 1994, Stanley, 1975, 1979), second only to  
speciation or origination; it can also be the ultimate manifestation of selection  
<sup>42</sup> (Jablonski, 2008, Rabosky and McCune, 2010).

While estimation of both trait-dependent speciation and extinction rates from  
<sup>44</sup> phylogenies of extant taxa has grown dramatically (Fitzjohn, 2010, Goldberg  
et al., 2011, 2005, Maddison et al., 2007, Rabosky et al., 2013, Stadler, 2011,  
<sup>46</sup> 2013, Stadler and Bokma, 2013), there are two major ways to estimate  
trait-dependent extinction: analysis of phylogenies, and analysis of the fossil  
<sup>48</sup> record. These two directions, phylogenetic comparative and paleobiological, are  
complementary and intertwined in the field of macroevolution (Hunt and  
<sup>50</sup> Rabosky, 2014, Jablonski, 2008, Rabosky and McCune, 2010). In the case of  
extinction, analysis of the fossil record has the distinct advantage over  
<sup>52</sup> phylogenies of only extant taxa because extinction is observable; this means  
that extinction rate is possible to estimate (Liow et al., 2010, Quental and  
<sup>54</sup> Marshall, 2009, Rabosky, 2010). The approach used here is thus complementary  
to the analysis of trait-dependent extinction based on a phylogeny.

<sup>56</sup> Jablonski (1986) observed that for bivalves at the end Cretaceous mass  
extinction event, the effects of some biological traits on taxonomic survival  
<sup>58</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>60</sup> multiple possible macroevolutionary mechanisms which may underlie this  
pattern: the effect of geographic range on survival remains constant and those of  
<sup>62</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

<sup>64</sup> traits decrease potentially by different degrees.

While Jablonski (1986) phrased his conclusions in terms of background versus  
<sup>66</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>68</sup> continuum (Payne and Finnegan, 2007, Simpson and Harnik, 2009, Wang, 2003).  
Additionally, in order to test the proposed macroevolutionary mechanism behind  
<sup>70</sup> the Jablonski (1986) scenario; not only do the taxon trait effects need to be  
modeled, but the correlation between trait effects need to be modeled as well.

<sup>72</sup> Here I model brachiopod taxon durations because trait based differences in  
extinction risk should manifest as differences in taxon durations. Brachiopods  
<sup>74</sup> 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  
<sup>76</sup> focus on the brachiopod record from the post-Cambrian Paleozoic, from the  
start of the Ordovician (approximately 485 My) through the end Permian  
<sup>78</sup> (approximately 252 My) as this represents the time of greatest global  
brachiopod diversity (Alroy, 2010).

<sup>80</sup> The analysis of taxon durations, or time from origination to extinction, falls  
under the purview of survival analysis, a field of applied statistics commonly  
<sup>82</sup> used in health care and engineering (Klein and Moeschberger, 2003) but has a  
long history in paleontology (Crampton et al., 2016, Simpson, 1944, 1953, Smits,  
<sup>84</sup> 2015, Van Valen, 1973, 1979). I adopt a hierarchical modeling approach  
(Gelman et al., 2013, Gelman and Hill, 2007, Klein and Moeschberger, 2003),  
<sup>86</sup> which represents both a conceptual and statistical unification of the  
paleontological dynamic and cohort survival analytic approaches (Baumiller,  
<sup>88</sup> 1993, Crampton et al., 2016, Ezard et al., 2012, Foote, 1988, Raup, 1975, 1978,  
Simpson, 2006, Smits, 2015, Van Valen, 1973, 1979).

<sup>90</sup> **1.1 Factors affecting brachiopod survival**

Conceptually, taxon survival can be considered an aspect of “taxon fitness”  
<sup>92</sup> along with expected lineage specific branching/origination rate (Cooper, 1984,  
Palmer and Feldman, 2012). Traits associated with taxon survival are thus  
<sup>94</sup> examples of species (or higher-level) selection, as differences in survival are  
analogous to differences in fitness. The traits analyzed here are all examples of  
<sup>96</sup> emergent and aggregate traits (Jablonski, 2008, Rabosky and McCune, 2010), in  
this case they are genus-level traits. Emergent traits are those which are not  
<sup>98</sup> measurable at a lower level (e.g. individual organism) such as geographic range,  
environmental affinity, or even fossil sampling rate. Aggregate traits, like body  
<sup>100</sup> size, are the average of a shared trait across all members of a lower level.

Geographic range is widely considered the most important biological trait for  
<sup>102</sup> estimating differences in extinction risk at nearly all times, with large  
geographic range associated with low extinction risk (Finnegan et al., 2012,  
<sup>104</sup> Harnik et al., 2012, Jablonski, 1986, 1987, 2008, Jablonski and Roy, 2003, Payne  
and Finnegan, 2007). This stands to reason even if extinction is completely at  
<sup>106</sup> random; a taxon with an unrestricted range is less likely to go extinct randomly  
than a taxon with a restricted range.

<sup>108</sup> 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  
<sup>110</sup> 100m. In contrast, open-ocean coastline environments have much greater  
variance in depth, do not cover the continental craton, and can persist during  
<sup>112</sup> 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  
<sup>114</sup> periods of low sea levels, such as during glacial periods, when epicontinental  
seas are drained. During the Paleozoic (approximately 541–252 My),

<sup>116</sup> epicontinental seas were widely spread globally but declined over the Mesozoic  
<sup>118</sup> (approximately 252–66 My) and have nearly disappeared during the Cenozoic  
<sup>120</sup> (approximately 66–0 My) as open-ocean coastlines became the dominant  
shallow-marine setting (Johnson, 1974, Miller and Foote, 2009, Peters, 2008).

<sup>122</sup> Miller and Foote (2009) demonstrated that during several mass extinctions taxa  
associated with open-ocean environments tend to have a greater extinction risk  
<sup>124</sup> than those taxa associated with epicontinental seas. During periods of  
background extinction, however, they found no consistent difference between  
<sup>126</sup> taxa favoring either environment. These two environment types represent the  
primary identifiable environmental dichotomy observed in ancient marine  
systems (Miller and Foote, 2009, Peters, 2008, Sheehan, 2001). Given these  
<sup>128</sup> findings, I predict that as extinction risk increases, the extinction risk associated  
with open-ocean environments should generally increase.

Because environmental preference is defined here as the continuum between  
<sup>130</sup> occurring exclusively in open-ocean environments versus epicontinental  
environments, intermediate values are considered “generalists” in the sense that  
<sup>132</sup> they favor neither end member. A long-standing hypothesis is that generalists or  
unspecialized taxa will have greater survival than specialists (Baumiller, 1993,  
<sup>134</sup> Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944). Because  
of this, the effect of environmental preference was modeled as a quadratic  
<sup>136</sup> function where a concave down relationship between preference and expected  
duration indicates that generalists are favored over specialists end-members.

<sup>138</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>140</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>142</sup> selectivity in Devonian brachiopods in both a phylogenetic and

non-phylogenetic context; finding that body size was not found to be associated  
144 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  
146 geographic range size in determining extinction risk (Harnik, 2011). Given these  
results, I expect that if body size has any effect on brachiopod taxonomic  
148 survival it is very small.

It is well known that, given the incompleteness of the fossil record, the observed  
150 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,  
152 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  
154 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  
156 relationship between the quality of sampling of a taxon and its apparent  
duration (e.g. greater sampling, longer duration). If there is no relationship  
158 between sampling and duration then interpretation can be made clearly; while  
observed durations are obviously truncated true durations, a lack of a  
160 relationship would indicate that the amount and form of this truncation is not a  
major determinant of the taxon's apparent duration.

## 162 **2 Materials and Methods**

### **2.1 Fossil occurrence information**

164 The brachiopod dataset analyzed here was sourced from the Paleobiology  
Database (<http://www.paleodb.org>) which was then filtered based on  
166 taxonomic (Rhyconelliformea: Rhynchonellata, Chileata, Obolellida,

Kutorginida, Strophomenida, Spiriferida)), temporal (post-Cambrian Paleozoic),  
168 stratigraphic, and other occurrence information used in this analysis. Analyzed  
occurrences were restricted to those with paleolatitude and paleolongitude  
170 coordinates, assignment to either epicontinental or open-ocean environment,  
and belonging to a genus present in the body size dataset (Payne et al., 2014).

172 Epicontinental versus open-ocean assignments for each fossil occurrence are  
partially based on those from Miller and Foote (2009), with additional  
174 occurrences assigned similarly (Miller and Foote, personal communication).

These filtering criteria are very similar to those from Foote and Miller (2013)  
176 with an additional constraint of being present in the body size data set from  
Payne et al. (2014).

178 Fossil occurrences were analyzed at the genus level which is common for  
paleobiological, macroevolutionary and macroecological studies of marine  
180 invertebrates (Alroy, 2010, Foote and Miller, 2013, Harnik et al., 2012, Kiessling  
and Aberhan, 2007, Miller and Foote, 2009, Nürnberg and Aberhan, 2013, 2015,  
182 Payne and Finnegan, 2007, Simpson and Harnik, 2009, Vilhena et al., 2013).

While species diversity dynamics are frequently of much greater interest than  
184 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  
186 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  
accuracy in the data analyzed here.

Genus duration was calculated as the number of geologic stages from first  
190 appearance to last appearance, inclusive. Durations were based on geologic  
stages as opposed to millions of years because of the inherently discrete nature  
192 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>194</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>196</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>198</sup> left censored (Klein and Moeschberger, 2003). How the likelihood of censored  
observations is calculated is detailed below.

<sup>200</sup> The covariates detailed below are geographic range size ( $r$ ), environmental  
preference ( $v, v^2$ ), body size ( $m$ ), and sampling ( $s$ ).

<sup>202</sup> Geographic range was calculated using an occupancy approach corrected for  
incomplete sampling. First, all occurrences were projected onto an equal-area  
<sup>204</sup> cylindrical map projection. Each occurrence was then assigned to one of the  
cells from a  $70 \times 34$  regular raster grid placed on the map. Each grid cell  
<sup>206</sup> represents approximately 250,000 km<sup>2</sup>. The map projection and regular lattice  
were made using shape files from <http://www.naturalearthdata.com/> and the  
<sup>208</sup> **raster** package for R (Hijmans, 2015). For each stage, the total number of  
occupied grid cells was calculated. Then, for each temporal bin, the relative  
<sup>210</sup> occurrence probability of the observed taxa was calculated using the JADE  
method developed by Chao et al. (2015). This method accounts for the fact that  
<sup>212</sup> taxa with an occupancy of 0 cannot be observed which means that occupancy is  
estimated from a truncated Binomial distribution. This correction is critical  
<sup>214</sup> when comparing occupancies from different times with different geographic  
sampling. Finally, for each genus, the mean relative occurrence probability was  
<sup>216</sup> calculated as the average of that genus' occurrence probabilities for all stages it  
was sampled.

Environmental preference was defined as probability of observing the ratio of  
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

at the same time ( $e'_i/E'_i$ ). This measure of environmental preference is expressed.

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

218 where  $v$  is the percent of the distribution defined in equation 2.1 less than or  
equal to  $\theta_i$ . The Beta distribution is used here because it is a continuous  
220 distribution bounded at 0 and 1, which is ideal for modeling percentages.

Body size, measured as shell length, was sourced directly from Payne et al.  
222 (2014). These measurements were made from brachiopod taxa figured in the  
*Treatise on Invertebrate Paleontology* (Williams et al., 2007).

224 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  
226 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  
228 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  
230 three stages. Sampling was only included as a covariate in the truncated model,  
described later.

232 Prior to analysis, geographic range was logit transformed and body size was  
natural-log transformed; both of these transformations make these variables  
234 defined for the entire real line. All covariates were then standardized by  
subtracting the mean from all values and dividing by twice its standard  
236 deviation, which follows Gelman and Hill (2007). This standardization means  
that all regression coefficients are comparable as the expected change per 1-unit  
238 change in the rescaled covariates.  $D$  is defined as the total number of covariate  
effects included in the model plus one for the intercept term.

<sup>240</sup> **2.2 Analytical approach**

Hierarchical modelling is a statistical approach which explicitly takes into 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 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 (e.g. prior) of all cohorts, observed and unobserved. The group-level parameters, 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. covariate effects) (Gelman et al., 2013). The subsequent estimates are partially pooled 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 overall group mean. All covariate effects (regression coefficients), as well as the intercept term (baseline extinction risk), were allowed to vary by group (origination cohort). The covariance between covariate effects was also modeled.

<sup>256</sup> Genus durations were assumed to follow a Weibull distribution which allows for age-dependent extinction (Klein and Moeschberger, 2003):  $y \sim \text{Weibull}(\alpha, \sigma)$ .

<sup>258</sup> The Weibull distribution has two parameters: scale  $\sigma$ , and shape  $\alpha$ . When  $\alpha = 1$ ,  $\sigma$  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 extinction risk increases with age, and values less than 1 indicate that extinction risk decreases with age. Note that the Weibull distribution is equivalent to the exponential distribution when  $\alpha = 1$ .

<sup>264</sup> In the case of the right- and left-censored observations mentioned above, the probability of those observations is calculated differently. For right-censored

<sup>266</sup> observations, the likelihood is calculated  $p(y|\theta) = 1 - F(y) = S(y)$  where  $F(y)$  is the cumulative distribution function (Klein and Moeschberger, 2003). In  
<sup>268</sup> contrast, the likelihood of a left-censored observation is calculated  $p(y|\theta) = F(y)$  (Klein and Moeschberger, 2003).

<sup>270</sup> The scale parameter  $\sigma$  was modeled as a regression following Kleinbaum and Klein (2005) with both varying intercept and varying slopes; this is expressed

$$\sigma_i = \exp\left(\frac{-\mathbf{X}_i B_{j[i]}}{\alpha_{j[i]}}\right) \quad (1)$$

<sup>272</sup> where  $i$  indexes across all observations,  $j[i]$  is the cohort membership of the  $i$ th observation,  $X$  is a  $N \times D$  matrix of covariates along with a column of 1's for  
<sup>274</sup> the intercept term, and  $B$  is a  $J \times D$  matrix of cohort-specific regression coefficients.

<sup>276</sup> Each of the rows of matrix  $B$  are modeled as realizations from a multivariate normal distribution with length  $D$  location vector  $\mu$  and  $J \times J$  covariance  
<sup>278</sup> matrix  $\Sigma$ :  $B_j \sim \text{MVN}(\mu, \Sigma)$ . The covariance matrix was then decomposed into a length  $J$  vector of scales  $\tau$  and a  $J \times J$  correlation matrix  $\Omega$ , defined  
<sup>280</sup>  $\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau)$  where “diag” indicates a diagonal matrix.

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

priors are defined

$$\mu_0 \sim \mathcal{N}(0, 5)$$

$$\mu_r \sim \mathcal{N}(-1, 1)$$

$$\mu_v \sim \mathcal{N}(0, 1)$$

$$\mu_{v^2} \sim \mathcal{N}(1, 1)$$

$$\mu_m \sim \mathcal{N}(0, 1)$$

$$\tau \sim C^+(1)$$

$$\Omega \sim LKJ(2).$$

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 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:

$$\alpha' \sim \mathcal{N}(0, 1)$$

$$\sigma_a \sim C^+(1)$$

$$\sigma_\alpha \sim C^+(1).$$

The above model is for all taxa and does not include sampling as a covariate. In  
 282 order to determine if sampling is acting as a confounding factor in this analysis,  
 an additional model was developed because sampling was only estimated for  
 284 taxa with a duration of three or more which creates a left-truncated distribution  
 of durations (Klein and Moeschberger, 2003). The sampling statement and  
 286 log-probability for a left-truncated Weibull distribution, truncated at time  $Y$

(e.g. three), is

$$\begin{aligned} p(y|\theta) &= \frac{\text{Weibull}(y, \alpha, \sigma)}{1 - \text{Weibull}_{cdf}(Y, \alpha, \sigma)} \\ &= \frac{\text{Weibull}(y, \alpha, \sigma)}{\text{Weibull}_{ccdf}(Y, \alpha, \sigma)} \end{aligned} \quad (2)$$

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

- 288 Note that cdf stands for cumulative distribution function and ccdf is the  
complementary cumulative distribution function.
- 290 The definition of  $\sigma$  (Eq. 1) is then updated so that  $X$ , the matrix of covariates,  
and  $B$ , the matrix of regression coefficients, now include an additional column  
292 for the sampling estimates and the cohort-specific effects of sampling. This  
addition then modifies the dimensions of  $\mu$  and  $\Sigma$ ; the new group-level effect of  
294  $\mu_s$  is given a weakly informative prior:  $\mu_s \sim \mathcal{N}(0, 1)$ .

For this left-truncated model, I've excluded one observation that is  
296 right-censored with a duration equal to the truncation time; the second line of  
equation 2 becomes  $p(y|\theta) = \text{Weibull}_{ccdf}(y, \alpha, \sigma)/\text{Weibull}_{ccdf}(Y, \alpha, \sigma)$  which  
298 yields a log-probability of 0.

The joint posterior was approximated using a Markov-chain Monte Carlo  
300 routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn  
Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic  
302 programming language Stan (Stan Development Team, 2014). The posterior  
distribution was approximated from four parallel chains run for 10,000 steps  
304 each, split half warm-up and half sampling and thinned to every 10th sample for  
a total of 4000 posterior samples. Chain convergence was assessed via the scale  
306 reduction factor  $\hat{R}$  where values close to 1 ( $\hat{R} < 1.1$ ) indicate approximate  
convergence. Convergence means that the chains are approximately stationary  
308 and the samples are well mixed (Gelman et al., 2013).

Model adequacy was evaluated using a couple of posterior predictive checks.

310 Posterior predictive checks are a means for understanding model fit or adequacy where the basic idea is that replicated data sets simulated from the fitted model  
312 should be similar to the original data and systematic differences between the simulations and observations indicate weaknesses of the model fit (Gelman  
314 et al., 2013). For both approaches used here, 1000 posterior predictive datasets were generated from 1000 unique draws from the posterior distribution of each  
316 parameter. The two posterior predictive checks used in this analysis are 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  
318 posterior predictive datasets, and comparison of the observed genus durations to the average posterior predictive estimate of  $\log(\sigma)$  (Eq. 1). The former is to see if simulated data has a similar survival pattern to the observed, while the latter  
320 is to see if the model systematically over- or under- estimates taxon survival.  
322

### 3 Results

324 Comparison of the posterior predictive estimates of  $S(t)$  to the empirical estimate reveal few obvious biases except for the case of estimates from the far  
326 right tail of observed durations (Fig. 1). This result is reinforced by the additional posterior predictive comparison where most estimates are not  
328 systematically biased except for a consistent under-estimate of  $\log(\sigma)$  for older taxa (Fig. 2). The results of both posterior predictive checks indicate that, for  
330 the majority of observations, model fit is generally not biased.

The cohort-level estimate of the effect of geographic range size indicates that as  
332 a taxon's geographic range increases, that taxon's duration is expected to increase (Table 1). Given the estimates of  $\mu_r$  and  $\tau_r$ , there is a less than 0.008%

<sup>334</sup> ( $\pm 0.05$  SD) probability that this relationships would be reversed  
 $(\Pr(\mathcal{N}(\mu_r, \tau_r) > 0))$ . The between-cohort variance  $\tau_r$  is the lowest of all the  
<sup>336</sup> regression coefficients (Table 1).

Body size is estimated to have no effect on taxon duration, with the estimate  
<sup>338</sup> being nearly 0 (Table 1). The variance between the cohort-specific estimates of  
the effect of body size  $\tau_m$  is estimated to be greater than the variance of  
<sup>340</sup> between-cohort estimates of the effect of geographic range size  $\tau_r$ .

The group-level estimate of the effect of environmental preference is estimated  
<sup>342</sup> from both  $\mu_v$  and  $\mu_{v^2}$ .

The estimate of  $\mu_v$  indicates that epicontinental favoring taxa are expected to  
<sup>344</sup> have a greater duration than open-ocean favoring taxa (Table 1). Additionally,  
given the estimate of between-cohort variance  $\tau_v$ , there is approximately 19%  
<sup>346</sup> ( $\pm 8$  SD) probability that, for any given cohort, taxa favoring open-ocean  
environments would have a greater expected duration than taxa favoring  
<sup>348</sup> epicontinental environments ( $\Pr(\mathcal{N}(\mu_v, \tau_v) > 0)$ ).

The estimate of  $\mu_{v^2}$  indicates that the overall relationship between  
<sup>350</sup> environmental preference and  $\log(\sigma)$  is concave down (Fig. 3), with only a 1.9%  
( $\pm 2.4$  SD) probability that any given cohort is convex up ( $\Pr(\mathcal{N}(\mu_{v^2}, \tau_{v^2}) < 0)$ ).

<sup>352</sup> The cohort-specific estimates of all the regression coefficients demonstrate a lot  
of between cohort variance, with no obvious trends. As indicated in Table 1 and  
<sup>354</sup> detectable visually (Fig. 4), the between-cohort estimates for  $\beta_0$ ,  $\beta_r$ , and  $\beta_m$  all  
have much lower variance than the between-cohort estimates of both  $\beta_v$  and  $\beta_{v^2}$ .

<sup>356</sup> While most cohort-specific estimates are very similar to the overall cohort-level  
estimate, there are a few notable excursions away from the overall mean (Fig.  
<sup>358</sup> 4). There are simultaneous excursions in both  $\beta_0$  and  $\beta_v$  for cohorts originating  
in the Givetian (387-382 My) and Frasnian (382-372 My) stages; both of which

<sup>360</sup> directly precede the late Devonian mass extinction event at the  
Frasnian/Famennian boundary. These cohorts are marked by both a high  
<sup>362</sup> extinction intensity and an increase in expected duration for taxa favoring  
epicontinental environments over open-ocean ones; this is consistent with the  
<sup>364</sup> results of Miller and Foote (2009).

Cohorts originating from the Silurian through the Early Devonian have a  
<sup>366</sup> slightly lower extinction intensity than the overall mean; these cohorts are those  
originating in the Llandovery (443-443 My) through the Emsian (407-393 My).  
<sup>368</sup> This is also a time period is also when there is the lowest overall probability  
that epicontinental favoring taxa are expected to have greater duration than  
<sup>370</sup> open-ocean favoring taxa. Both the Silurian and Devonian periods are notable  
for having been periods with a mostly “hothouse” climate, with no polar icecaps  
<sup>372</sup> and a high sea-level (Edwards and Fanning, 1985, Joachimski et al., 2009,  
Munnecke et al., 2010).

<sup>374</sup> The cohort-specific relationships between environmental preference and  $\log(\sigma)$   
were calculated from the estimates of  $\beta_0$ ,  $\beta_v$ , and  $\beta_{v^2}$  (Fig. 5) and reflect how  
<sup>376</sup> 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  
<sup>378</sup> is notable that the cohort originating in the Kungurian (279-272 My) is least  
like the overall expected relationship and has the most sharply curved  
<sup>380</sup> appearance due to a high estimate  $\beta_{v^2}$  (Fig. 4). This cohort has the biggest  
difference in extinction risk between environmental generalists and specialists.  
<sup>382</sup> The cohorts originating during the Emsian (407-393 My) and Frasnian (382 -  
372 My) are tied for second in sharpness of curvature. The least sharply curved  
<sup>384</sup> cohorts include those originating during Tremadocian (484-477 My), Hirnantian  
(445-443 My), Llandovery (443-433 My), and Ludlow (427-423 My). Except for  
<sup>386</sup> the Tremadocian cohort, most of these cohorts originate during the Silurian

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

The correlations of the cohort-specific estimates of the regression coefficients are  
390 estimated as the off-diagonal elements of the correlation matrix  $\Omega$ . Only two of  
the elements of  $\Omega$  are distinguishable from 0: the correlation of  $\beta_0$  (extinction  
392 intensity) with both  $\beta_r$  and  $\beta_v$  (Fig. 6).

There is an approximate 86% probability that the cohort-specific estimates of  
394 baseline extinction intensity  $\beta_0$  and the effect of geographic range  $\beta_r$  are  
negatively correlated; this means that for cohorts experiencing a lower  
396 extinction intensity ( $\beta_0$  decreases), the magnitude of the effect of geographic  
range is expected to decrease as well, and *vice versa*; this is in contrast to the  
398 observation made by Jablonski (1986) with regards to late Cretaceous bivalves.

Similarly, there is an approximate 99.9% probability that the cohort-specific  
400 estimates of  $\beta_0$  and  $\beta_v$  are negatively correlated; this means that as extinction  
intensity increases it is expected that epicontinental taxa become more favored  
402 over open-ocean environments (i.e. as  $\beta_0$  increases,  $\beta_v$  decreases). Additionally,  
there is only a 19% ( $\pm 8.1$  SD) probability that, for any given cohort,  
404 open-ocean environments will be favored ( $\Pr(|\mu_v| - 2\tau_v > 0)$ ).

There is only an approximate 68% probability that  $\beta_r$  and  $\beta_v$  are positively  
406 correlated. This lack of cross-correlation may be due in part to the much higher  
between-cohort variance of the effect of environmental preference  $\tau_v$  than the  
408 very small between-cohort variance in the effect of geographic range  $\tau_r$  (Table  
1); the effect of geographic range might simply not vary enough relative to the  
410 much noisier environmental preference.

## 4 Discussion

<sup>412</sup> The generating observation behind this study was that for bivalves at the end  
<sup>413</sup> Cretaceous mass extinction event, the only biological trait that was found to  
<sup>414</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>416</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>418</sup> with changing extinction intensity?

I find that as intensity increases ( $\beta_0$  decreases), the magnitude of the effect of  
<sup>420</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>422</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>424</sup> environmental preference. However, the between-cohort variance in effect of  
geographic range is much less than the between-cohort variance of the effect of  
<sup>426</sup> environmental preference. As stated earlier, this may underlie the lack of  
correlation between these two effects.

<sup>428</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>430</sup> greater than average extinction intensity, the difference in the effect of geographic  
range and the group-level effect of geographic range is expected to be smaller  
<sup>432</sup> than the difference between the effect of environmental preference and the  
group-level effect of environmental preference.

<sup>434</sup> I find consistent support for the “survival of the unspecialized,” with respect to  
epicontinental versus open-ocean environmental preference, as a time-invariant  
<sup>436</sup> generalization of brachiopod survival; taxa with intermediate environmental

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

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

An alternative approach for specifically modeling survival that can take into  
account imperfect observation than the method used here is the  
Cormack-Jolly-Seber (CJS) model (Liow et al., 2008, Liow and Nichols, 2010,

<sup>464</sup> Royle and Dorazio, 2008, Tomiya, 2013). This model is a type of hidden Markov  
<sup>465</sup> model with an absorbing state (i.e. extinction). In this model, survival is defined  
<sup>466</sup> as the probability of surviving from time  $t$  to time  $t + 1$ . Additionally, the effect  
<sup>467</sup> of preservation and sighting is estimated as probability of observing a taxon  
<sup>468</sup> that is present; this can extend the duration of a taxon beyond its last  
<sup>469</sup> occurrence. This approach is a fundamentally different from the method used in  
<sup>470</sup> the above analysis; for example, instead of modeling persistence from time  $t$  to  
<sup>471</sup> time  $t + 1$ , I model the distribution of taxon durations (lifetime) which then  
<sup>472</sup> decomposes into a function describing survival ( $S(t)$ ) and a function describing  
<sup>473</sup> instantaneous extinction rate with taxon age (Klein and Moeschberger, 2003,  
<sup>474</sup> Royle and Dorazio, 2008).

<sup>475</sup> The use of genera as the unit of the study and how to exactly interpret the  
<sup>476</sup> effects of the biological traits is an important question. For example, if any of  
<sup>477</sup> the traits analyzed here are associated with increases in speciation rates, this  
<sup>478</sup> might increase the duration of genera through self-renewal (Raup, 1991, 1994),  
<sup>479</sup> which would be an example of the difference in biological pattern between  
<sup>480</sup> species and genera (Jablonski, 1987, 2007, 2008). This could lead to a trait  
<sup>481</sup> appearing to decrease generic level extinction risk by that trait increasing  
<sup>482</sup> species level origination rate instead of decreasing species level extinction risk.

<sup>483</sup> The model used here could be improved through either increasing the number of  
<sup>484</sup> analyzed traits, expanding the hierarchical structure of the model to include  
<sup>485</sup> other major taxonomic groups of interest, and the inclusion of explicit  
<sup>486</sup> phylogenetic relationships between the taxa in the model as an additional  
<sup>487</sup> hierarchical effect. An example trait that may be of particular interest is the  
<sup>488</sup> affixing strategy or method of interaction with the substrate of the taxon, which  
<sup>489</sup> has been found to be related to brachiopod survival where, for cosmopolitan  
<sup>490</sup> taxa, taxa that are attached to the substrate are expected to have a greater

duration than those that are not (Alexander, 1977).

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

Traits like environmental preference or geographic range (Hunt et al., 2005,  
500 Jablonski, 1987) are most likely heritable. Without phylogenetic context, this  
analysis assumes that differences in extinction risk between taxa are independent  
502 of the shared evolutionary history of those taxa (Felsenstein, 1985). In contrast,  
the origination cohorts only capture shared temporal context. For example, if  
504 taxon duration is phylogenetically heritable, then closely related taxa may have  
more similar durations as well as more similar biological traits. Without taking  
506 into account phylogenetic similarity the effects of these biological traits would  
be inflated solely due to inheritance. The inclusion of phylogenetic context as an  
508 additional individual-level hierarchical effect independent of origination cohort  
would allow for determining how much of the observed variability is due to  
510 shared evolutionary history versus shared temporal context versus actual  
differences associated with biological traits (Harnik et al., 2014, Smits, 2015).

512 In summary, patterns of Paleozoic brachiopod survival were analyzed using a  
fully Bayesian hierarchical survival modelling approach while also eschewing the  
514 traditional separation between background and mass extinction. I find that  
cohort extinction intensity is negatively correlated with both the cohort-specific  
516 effects of geographic range and environmental preference. These results imply  
that as extinction intensity increases ( $\beta_0$ ) increases, it is expected that both

518 effects will increase in magnitude. However, the change in effect of  
environmental preference is expected to be greater than the change in the effect  
520 of geographic range. Additionally, I find support for greater survival in  
environmental generalists over specialists in all origination cohorts analyzed;  
522 this is consistent with the long standing “survival of the unspecialized”  
hypothesis (Baumiller, 1993, Liow, 2004, 2007, Nürnberg and Aberhan, 2013,  
524 2015, Simpson, 1944, 1953, Smits, 2015). I also find that taxa favoring  
epicontinental environments are expected to have a slightly greater duration  
526 than open-ocean favoring taxa. The results of this analysis support the  
conclusion that for Paleozoic brachiopods, as extinction intensity increases  
528 overall extinction selectivity is expected to increase as well.

## Acknowledgements

530 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 thank  
532 D. Bapst, N. Pierrehumbert and M. Villarosa Garcia for additional comments.  
An additional thank you to A. Miller for the epicontinental versus open-ocean  
534 assignments. Finally, thank you to the reviewers for their helpful comments that  
improved this manuscript. This entire study would not have been possible  
536 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,  
538 M. Clapham, F. Fürsich, N. Heim, A. Hendy, S. Holland, L. Ivany, W. Kiessling,  
B. Kröger, A. McGowan, T. Olszewski, P. Novack-Gottshall, M. Patzkowsky, M.  
540 Uhen, L. Villier, and P. Wager. This work was supported by a NASA  
Exobiology grant (NNX10AQ446) to A. Miller and M. Foote. I declare no  
542 conflicts of interest. This is Paleobiology Database publication XXX.

## References

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

- 568 Felsenstein, J., 1985. Phylogenies and the comparative method. *American  
Naturalist* 125:1–15.
- 570 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.  
572 *Proceedings of the National Academy of Sciences* 109:6829–6834.
- 574 Fitzjohn, R. G., 2010. Quantitative Traits and Diversification. *Systematic  
Biology* Pp. 1–15.
- Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites.  
576 *Paleobiology* 14:258–271.
- , 2000a. Origination and extinction components of taxonomic diversity:  
578 general problems. *Paleobiology* 26:74–102.
- , 2000b. Origination and extinction components of taxonomic diversity:  
580 Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- , 2014. Environmental controls on geographic range size in marine animal  
582 genera. *Paleobiology* 40:440–458.
- Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine  
584 animal genera. *Paleobiology* 39:171–192.
- Foote, M. and D. Raup, 1996. Fossil preservation and the stratigraphic ranges  
586 of taxa. *Paleobiology* 22:121–140.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B.  
588 Rubin, 2013. *Bayesian data analysis*. 3 ed. Chapman and Hall, Boca Raton,  
FL.
- 590 Gelman, A. and J. Hill, 2007. *Data Analysis using Regression and  
Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY.

- 592 Goldberg, E. E., L. T. Lancaster, and R. H. Ree, 2011. Phylogenetic inference  
of reciprocal effects between geographic range evolution and diversification.
- 594 Systematic biology 60:451–65.
- Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski, 2005. Diversity,  
596 endemism, and age distributions in macroevolutionary sources and sinks. The  
American naturalist 165:623–33.
- 598 Harnik, P. G., 2011. Direct and indirect effects of biological factors on  
extinction risk in fossil bivalves. Proceedings of the National Academy of  
600 Sciences of the United States of America 108:13594–9.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson, 2014.  
602 Phylogenetic signal in extinction selectivity in Devonian terebratulide  
brachiopods. Paleobiology 40:675–692.
- 604 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  
606 Society B: Biological Sciences 279:4969–4976.
- Heim, N. A. and S. E. Peters, 2011. Regional environmental breadth predicts  
608 geographic range and longevity in fossil marine genera. PloS one 6:e18946.
- Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL  
610 <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  
612 natural selection on traits above the species level. Methods in Ecology and  
Evolution .
- 614 Hoffman, M. D. and A. Gelman, 2014. The No-U-Turn Sampler: Adaptively  
Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine  
616 Learning Research 15:1351–1381.

- Hopkins, M. J., C. Simpson, and W. Kiessling, 2014. Differential niche dynamics  
618 among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- Hunt, G. and D. L. Rabosky, 2014. Phenotypic Evolution in Fossil Species:  
620 Pattern and Process. *Annual Review of Earth and Planetary Sciences*  
42:421–441.
- Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a  
622 comment on "On the heritability of geographic range sizes". *American  
Naturalist* 166:129–135.
- Jablonski, D., 1986. Background and mass extinctions: the alternation of  
626 macroevolutionary regimes. *Science* 231:129–133.
- , 1987. Heritability at the species level: analysis of geographic ranges of  
628 cretaceous mollusks. *Science* 238:360–363.
- , 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- 630 ———, 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  
632 living molluscs. *Proceedings. Biological sciences / The Royal Society*  
634 270:401–6.
- Joachimski, M. M., S. Breisig, W. Buggisch, J. A. Talent, R. Mawson,  
636 M. Gereke, J. R. Morrow, J. Day, and K. Weddige, 2009. Devonian climate  
and reef evolution: Insights from oxygen isotopes in apatite. *Earth and  
638 Planetary Science Letters* 284:599–609.
- Johnson, J. G., 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- 640 Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine

- benthic biodiversity dynamics through Triassic–Jurassic time. *Paleobiology*  
642 33:414–434.
- Klein, J. P. and M. L. Moeschberger, 2003. Survival Analysis: Techniques for  
644 Censored and Truncated Data. 2nd ed. Springer, New York.
- Kleinbaum, D. G. and M. Klein, 2005. Survival analysis: a self-learning text. 2  
646 ed. Springer, New York, NY.
- Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively  
648 unspecialized" using fossil crinoids. *The American Naturalist* 164:431–43.
- , 2007. Does versatility as measured by geographic range, bathymetric  
650 range and morphological variability contribute to taxon longevity? *Global  
Ecology and Biogeography* 16:117–128.
- 652 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  
654 mammals. *Proceedings of the National Academy of Sciences* 105:6097.
- Liow, L. H. and J. D. Nichols, 2010. Estimating rates and probabilities of  
656 origination and extinction using taxonomic occurrence data:  
Capture-mark-recapture (CMR) approaches. *Quantitative Methods in  
658 Paleobiology* Pp. 81–94.
- Liow, L. H., T. B. Quental, and C. R. Marshall, 2010. When can decreasing  
660 diversification rates be detected with molecular phylogenies and the fossil  
record? *Systematic Biology* 59:646–59.
- Maddison, W. P., P. E. Midford, and S. P. Otto, 2007. Estimating a binary  
662 character's effect on speciation and extinction. *Systematic Biology* 56:701.
- Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings:  
664 the kinetics of mass extinction and origination. *Science* 326:1106–9.

- 666 Munnecke, A., M. Calner, D. A. T. Harper, and T. Servais, 2010. Ordovician  
and Silurian sea-water chemistry, sea level, and climate: A synopsis.
- 668 Palaeogeography, Palaeoclimatology, Palaeoecology 296:389–413.
- Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range  
670 predict diversity dynamics in marine Mesozoic bivalves. Paleobiology  
39:360–372.
- 672 ———, 2015. Interdependence of specialization and biodiversity in Phanerozoic  
marine invertebrates. Nature communications 6:6602.
- 674 Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental  
than evolvability. PloS one 7:e38025.
- 676 Payne, J. L. and S. Finnegan, 2007. The effect of geographic range on  
extinction risk during background and mass extinction. Proceedings of the  
678 National Academy of Sciences 104:10506–11.
- Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain, 2014. Metabolic  
680 dominance of bivalves predates brachiopod diversity decline by more than 150  
million years. Proceedings of the Royal Society B 281:20133122.
- 682 Peters, S. E., 2008. Environmental determinants of extinction selectivity in the  
fossil record. Nature 454:626–9.
- 684 Quental, T. B. and C. R. Marshall, 2009. Extinction during evolutionary  
radiations: reconciling the fossil record with molecular phylogenies. Evolution;  
686 international journal of organic evolution 63:3158–67.
- Rabosky, D. L., 2010. Extinction rates should not be estimated from molecular  
688 phylogenies. Evolution 64:1816–24.
- Rabosky, D. L. and A. R. McCune, 2010. Reinventing species selection with  
690 molecular phylogenies. Trends in ecology & evolution 25:68–74.

- Rabosky, D. L., F. Santini, J. Eastman, S. a. Smith, B. Sidlauskas, J. Chang,  
692 and M. E. Alfaro, 2013. Rates of speciation and morphological evolution are  
correlated across the largest vertebrate radiation. *Nature Communications*  
694 4:1–8.
- Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law.  
696 *Paleobiology* 1:82–96.
- , 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- 698 ———, 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- , 1994. The role of extinction in evolution. *Proceedings of the National  
700 Academy of Sciences* 91:6758–6763.
- Royle, J. A. and R. M. Dorazio, 2008. Hierarchical modeling and inference in  
702 ecology: the analysis of data from populations, metapopulations and  
communities. Elsevier, London.
- Sepkoski, J. J., 1981. A factor analytic description of the Phanerozoic marine  
704 fossil record. *Paleobiology* 7:36–53.
- Sheehan, P., 2001. The late Ordovician mass extinction. *Annual Review of  
706 Earth and Planetary Sciences* 29:331–364.
- 708 Simpson, C., 2006. Levels of selection and large-scale morphological trends.  
Ph.D. thesis, University of Chicago.
- Simpson, C. and P. G. Harnik, 2009. Assessing the role of abundance in marine  
710 bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.
- 712 Simpson, G. G., 1944. *Tempo and Mode in Evolution*. Columbia University  
Press, New York.

- <sup>714</sup> ———, 1953. *The Major Features of Evolution*. Columbia University Press, New York.
- <sup>716</sup> 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.
- <sup>720</sup> Solow, A. and W. Smith, 1997. On fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 23:271–277.
- <sup>722</sup> Stadler, T., 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences* Pp. 1–6.
- <sup>724</sup> ———, 2013. Recovering speciation and extinction dynamics based on phylogenies. *Journal of evolutionary biology* 26:1203–1219.
- <sup>726</sup> Stadler, T. and F. Bokma, 2013. Estimating speciation and extinction rates for phylogenies of higher taxa. *Systematic biology* 62:220–30.
- <sup>728</sup> Stan Development Team, 2014. Stan: A c++ library for probability and sampling, version 2.5.0. URL <http://mc-stan.org/>.
- <sup>730</sup> Stanley, S. M., 1975. A theory of evolution above the species level. *Proc. Natl. Acad. Sci. USA* 72:646.
- <sup>732</sup> ———, 1979. *Macroevolution: pattern and process*. W. H. Freeman, San Francisco.
- Team, S. D., 2016. Stan Modeling Language Users Guide and Reference Manual.
- <sup>734</sup> Tomiya, S., 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist* 182:196–214.
- <sup>736</sup> Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- , 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.

- 738 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  
740 latitudinal selectivity gradient at the end-Cretaceous mass extinction.  
Scientific Reports 3:1790.
- 742 Wagner, P. J. and J. D. Marcot, 2013. Modelling distributions of fossil sampling  
rates over time, space and taxa: assessment and implications for  
744 macroevolutionary studies. Methods in Ecology and Evolution 4:703–713.
- Wang, S. C., 2003. On the continuity of background and mass extinction.  
746 Paleobiology 29:455–467.
- Wang, S. C. and C. R. Marshall, 2004. Improved confidence intervals for  
748 estimating the position of a mass extinction boundary. Paleobiology 30:5.
- Williams, C., C. H. C. Brunton, and S. J. Carlson, 2007. Treatise on  
750 invertebrate paleontology. Part H, Brachiopoda. Geological Society of  
America, Boulder, Colorado.

Table 1: 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 (SD), 10th, 50th, and 90th quantiles of the posterior are presented.

parameter	interpretation	mean	SD	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 <sup>2</sup>	2.88	0.31	2.48	2.88	3.27
$\mu_m$	body size	0.04	0.12	-0.12	0.04	0.19
$\tau_i$	SD(intercept)	0.50	0.10	0.37	0.49	0.63
$\tau_r$	SD(geographic range)	0.27	0.13	0.11	0.26	0.45
$\tau_v$	SD(environmental preference)	0.76	0.16	0.56	0.74	0.97
$\tau_{v^2}$	SD(environmental preference <sup>2</sup> )	1.24	0.33	0.84	1.21	1.67
$\tau_m$	SD(body size)	0.47	0.12	0.33	0.47	0.63

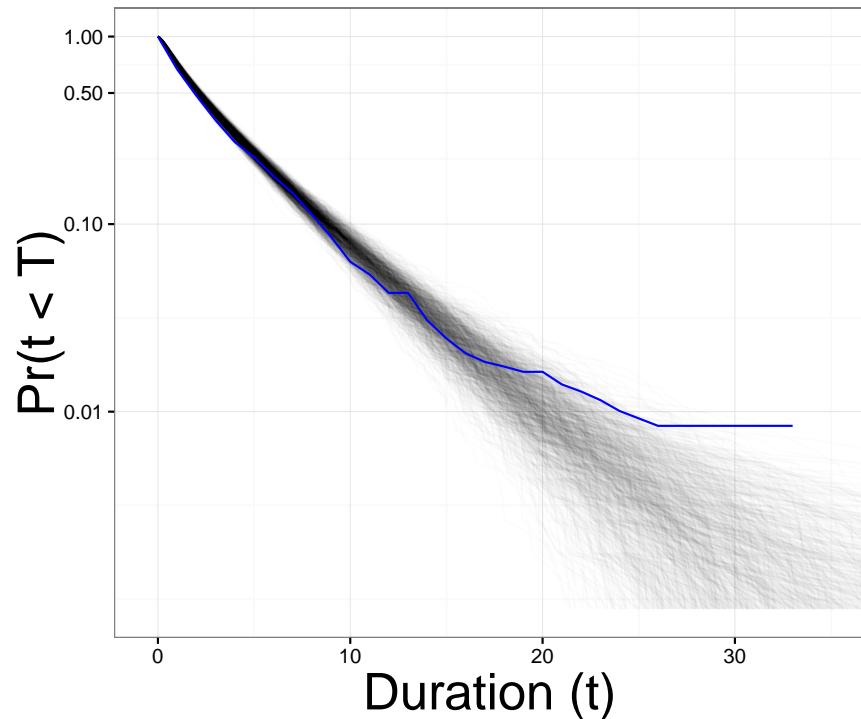


Figure 1: Comparison of the empirical estimate of  $S(t)$  (highlighted) versus estimates from 1000 posterior predictive data sets (black).  $S(t)$  corresponds to the probability that the age of a genus  $t$  is less than the genus' ultimate duration  $T$ . The vertical axis is log10 transformed.

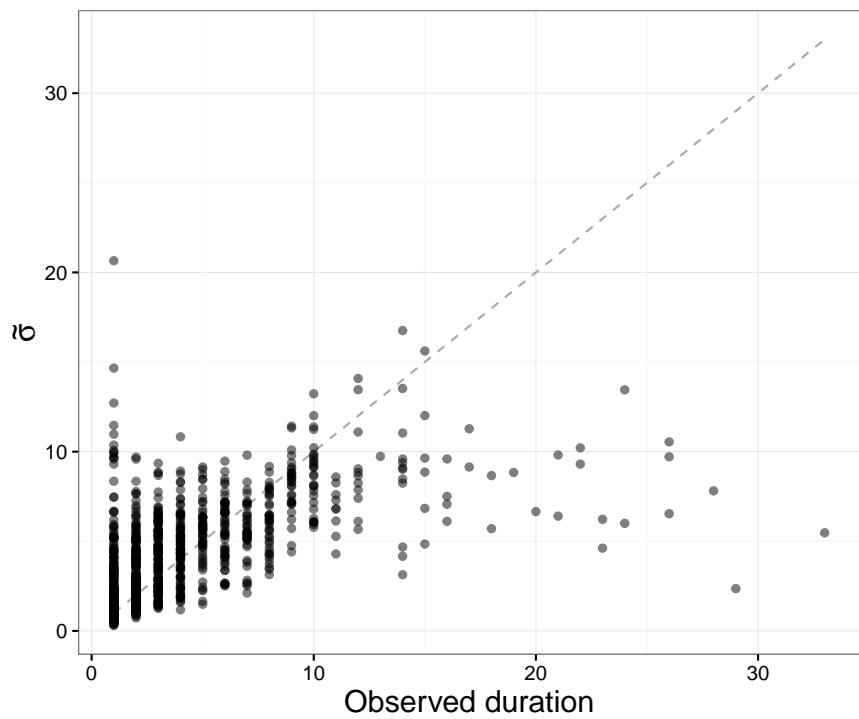


Figure 2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of  $\log(\sigma)$ . The dashed, diagonal line corresponds to  $x = y$ .

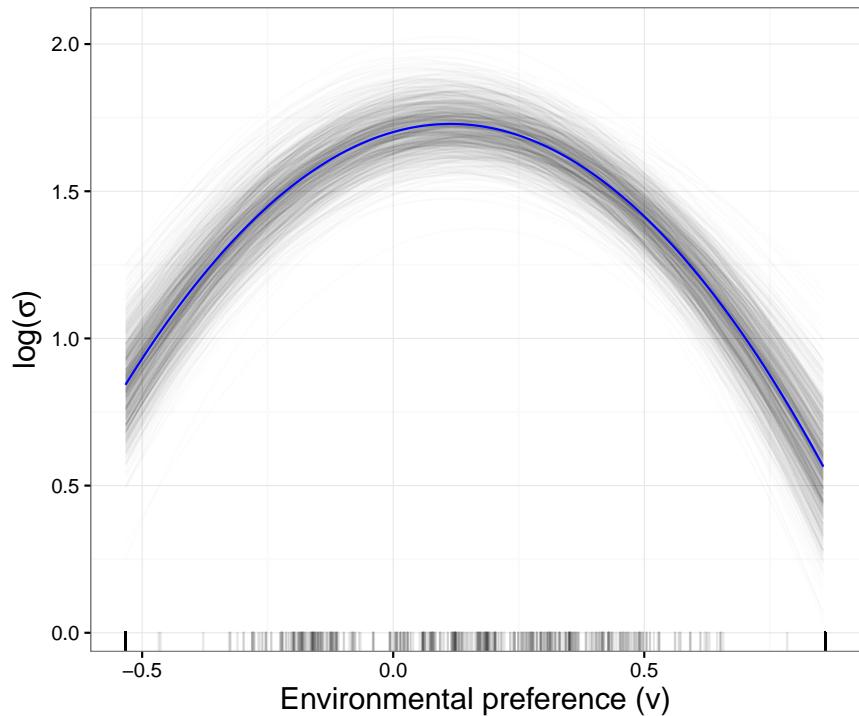


Figure 3: The overall expected relationship between environmental affinity  $v_i$  and a  $\log(\sigma)$  when  $r = 0$  and  $m = 0$ . Each grey line corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive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

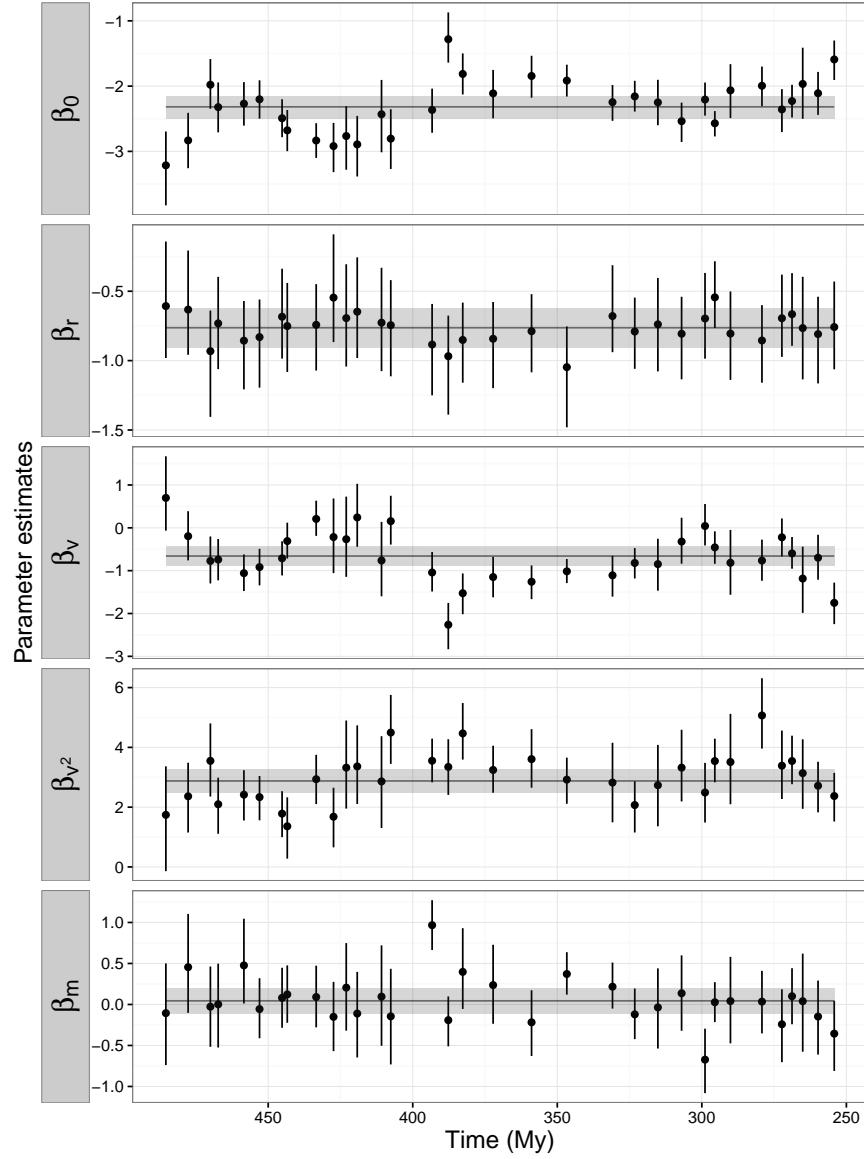


Figure 4: Comparison of cohort-specific estimates of  $\beta_0$ , the effect of geographic range on extinction risk  $\beta_r$ , the effect of environmental preference  $\beta_v$  and  $\beta_{v^2}$ , and body size  $\beta_m$ . Points correspond to the median of the cohort-specific estimate, along with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

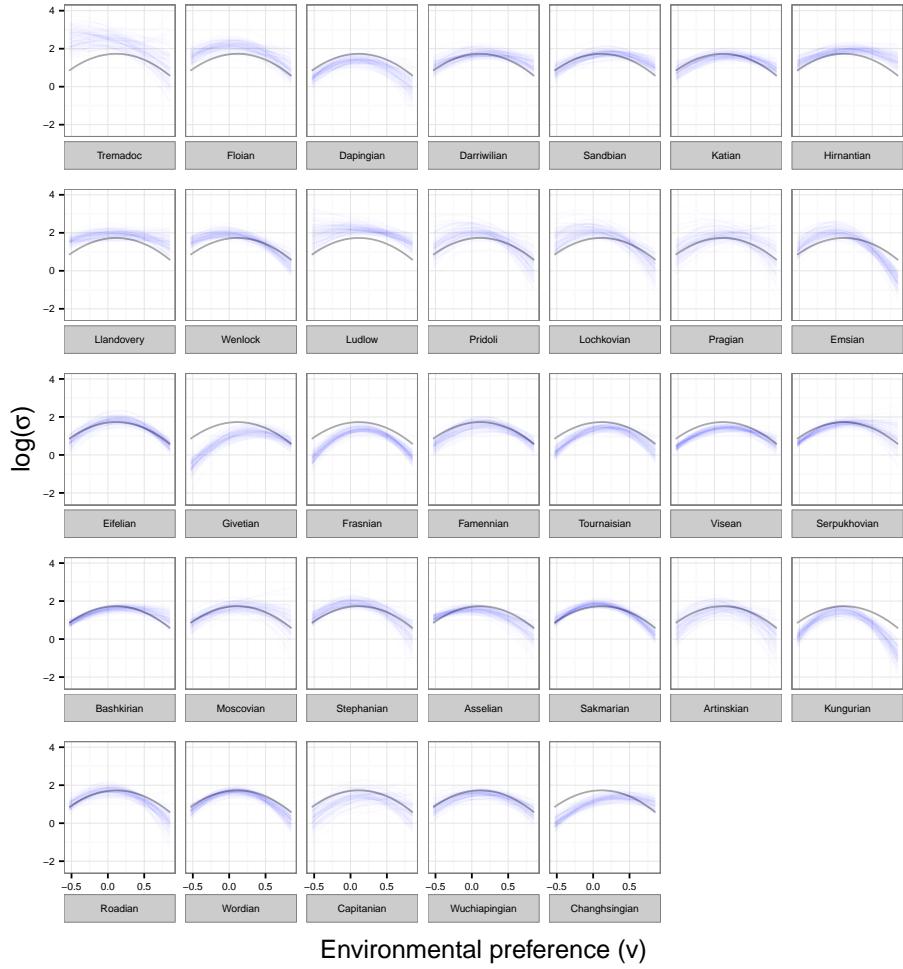


Figure 5: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on  $\log(\sigma)$  to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the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

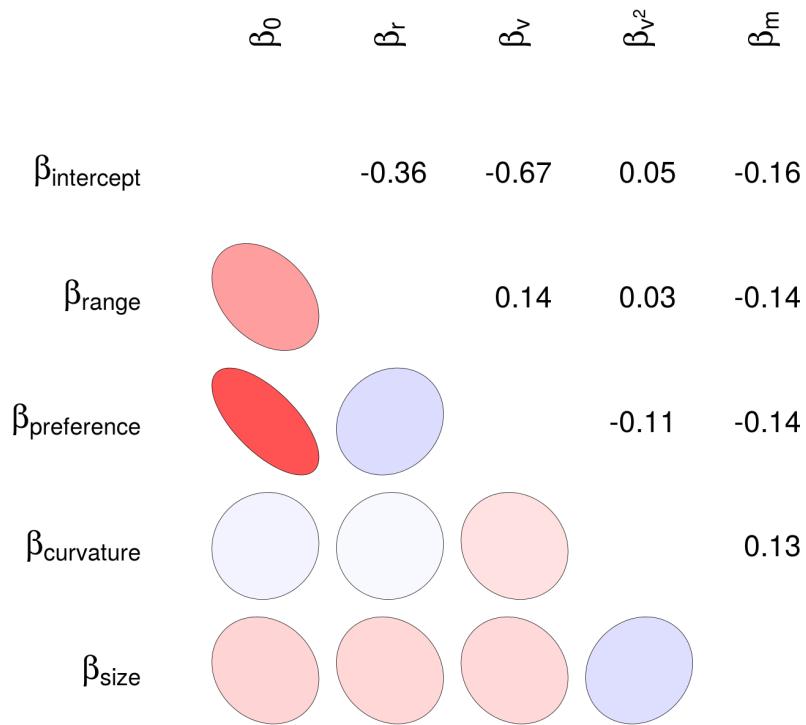


Figure 6: Mixed graphical and numerical representation of the correlation matrix  $\Omega$  of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates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