

**Title:** The interplay between extinction intensity and selectivity:  
2 correlation in trait effects on taxonomic survival

**Running title:** Variation in trait effects on taxonomic survival

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

6 **Keywords:** macroevolution, extinction, macroecology, Bayesian, brachiopods

**Word count:** approximately 5800.

8 **Table count:** 1.

**Figure count:** 5.

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

## Abstract

While the effect of geographic range on extinction risk is well documented, how other traits may increase or decrease extinction risk is less well known. 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 allow for environmental affinity to have a nonlinear effect on duration. Using a hierarchical Bayesian modeling approach, I also model the interaction between the effects of the biological traits and a taxon's time of origination. My analysis framework eschews the traditional distinction between background and mass extinction, instead the entire time period is analyzed where these "states" are part of a continuum. For parts of the Paleozoic I find support for a "survival of the generalists" scenario, though there are times where this relationship is absent or even reversed. Importantly, I find evidence that as baseline extinction risk increases, the size of the effect of geographic range increases but the effect of environmental preference tends to decrease. Additionally, I find strong evidence for correlation between the effects of geographic range and the non-linearity aspect of environmental preference which may help explain this pattern; taxa with large geographic ranges encompass more possible environments, so that when extinction intensity is high there is no difference in the environmental preference of the surviving taxa.

## 1 Introduction

How do biological traits affect extinction risk? Jablonski (1986) observed that at a mass extinction event, the effects of biological traits on taxonomic survival decreased in size. However, this pattern was not the case for the effect of geographic range on survival (Jablonski, 1986). Biological traits are defined here

as descriptors of a taxon's adaptive zone, which is the set of all biotic–biotic and  
40 biotic–abiotic interactions that a taxon can experience (Simpson, 1944). In effect, these are descriptors of a taxon's broad-sense ecology.

42 Jablonski (1986) phrased his conclusions in terms of background versus mass extinction, but this scenario is readily transferable to a continuous variation  
44 framework as there is no obvious distinction in terms of extinction rate between these two states (Wang, 2003). Additionally, the Jablonski (1986) scenario has  
46 strong model structure requirements in order to test its proposed macroevolutionary mechanism; not only do the taxon trait effects need to be modeled, but the correlation between trait effects need to be modeled as well.  
48

There are two end-member macroevolutionary mechanisms which may underlie  
50 the pattern observed by Jablonski (1986): the effect of geographic range on predictive survival remains constant and those of other biological traits decrease,  
52 and the effect of geographic range in predicting survival increases and those of other biological traits stay constant. Reality, of course, may fall somewhere  
54 along this continuum.

I choose to model brachiopod taxon durations because trait based differences in  
56 extinction risk should manifest as differences in taxon durations. Namely, a taxon with a beneficial trait should persist for longer, on average, than a taxon  
58 without that beneficial trait. Conceptually, taxon survival can be considered an aspect of “taxon fitness” along with expected lineage specific  
60 branching/origin rate (Cooper, 1984, Palmer and Feldman, 2012).

Brachiopods are an ideal group for this study as they are well known for  
62 having an exceptionally complete fossil record (Foote, 2000). Specifically, I focus on the brachiopod record from most of the Paleozoic, from the start of the  
64 Ordovician (approximately 485 Mya) through the end Permian (approximately 252 Mya) as this represents the time of greatest global brachiopod diversity

<sup>66</sup> (Alroy, 2010).

The analysis of taxon durations, or time from origination to extinction, falls  
<sup>68</sup> 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  
<sup>70</sup> paleontology (Simpson, 1944, 1953, Van Valen, 1973, 1979).

Geographic range is widely considered the most important taxon trait for  
<sup>72</sup> estimating differences in extinction risk at nearly all times with large geographic  
range associated with low extinction risk (Jablonski, 1986, 1987, Jablonski and  
<sup>74</sup> Roy, 2003, Payne and Finnegan, 2007), though Foote and Miller (2013) find  
that this generalization does not hold in the Mesozoic. For the Paleozoic,  
<sup>76</sup> however, I expect this to hold true for the entire period analyzed.

Miller and Foote (2009) demonstrated that during several mass extinctions taxa  
<sup>78</sup> associated with open-ocean environments tend to have a greater extinction risk  
than those taxa associated with epicontinental seas. During periods of  
<sup>80</sup> background extinction, however, they found no consistent difference between  
taxa favoring either environment. These two environment types represent the  
<sup>82</sup> primary environmental dichotomy observed in ancient marine systems (Miller  
and Foote, 2009, Peters, 2008, Sheehan, 2001).

<sup>84</sup> Epicontinental seas are a shallow-marine environment where the ocean has  
spread over the surface of a continental shelf with a depth typically less than  
<sup>86</sup> 100m. In contrast, open-ocean coastline environments have much greater  
variance in depth, do not cover the continental shelf, and can persist during  
<sup>88</sup> periods of low sea level. Because of this, it is strongly expected that taxa which  
favor epicontinental seas would be at great risk during periods of low sea levels,  
<sup>90</sup> such as during glacial periods, where these seas are drained. During the  
Paleozoic (approximately 541–252 My), epicontinental seas were widely spread

<sup>92</sup> globally but declined over the Mesozoic (approximately 252–66 My) and  
<sup>93</sup> eventually nearly disappeared during the Cenozoic (approximately 66–0 My) as  
<sup>94</sup> open-ocean coastlines became the dominant shallow-marine setting (Johnson,  
1974, Miller and Foote, 2009, Peters, 2008).

<sup>96</sup> Given the above findings, I predict that as extinction risk increases, taxa  
<sup>97</sup> associated with open-ocean environments should generally increase in extinction  
<sup>98</sup> risk versus those that favor epicontinental seas. Additionally, there is a possible  
<sup>99</sup> nonlinear relationship between environmental preference and taxon duration. A  
<sup>100</sup> long standing hypothesis is that generalists or unspecialized taxa will have  
<sup>101</sup> greater survival than specialists (Baumiller, 1993, Liow, 2004, 2007, Nürnberg  
<sup>102</sup> and Aberhan, 2013, 2015, Simpson, 1944). In this analysis I allowed for  
environmental preference to have a parabolic effect on taxon duration

<sup>104</sup> Body size, measured as shell length (Payne et al., 2014), was also considered as  
<sup>105</sup> a potentially informative covariate. Body size is a proxy for metabolic activity  
<sup>106</sup> and other correlated life history traits (Payne et al., 2014). There is no strong  
<sup>107</sup> hypothesis of how body size effects extinction risk in brachiopods, such that a  
<sup>108</sup> positive, negative, or zero effect are all plausible.

I adopt a hierarchical Bayesian survival modeling approach, which represents  
<sup>110</sup> both a conceptual and statistical unification of the paleontological dynamic and  
cohort survival analytic approaches (Baumiller, 1993, Foote, 1988, Raup, 1975,  
1978, Simpson, 2006, Van Valen, 1973, 1979). By using a Bayesian framework I  
<sup>112</sup> am able to quantify the uncertainty inherent in the estimates of the effects of  
<sup>113</sup> biological traits on survival, especially in cases where the covariates of interest  
<sup>114</sup> (i.e. biological traits) are themselves known with error.

<sub>116</sub> **2 Materials and Methods**

**2.1 Fossil occurrence information**

<sub>118</sub> The dataset analyzed here was sourced from the Paleobiology Database  
(<http://www.paleodb.org>) which was then filtered based on taxonomic,  
<sub>120</sub> temporal, stratigraphic, and other occurrence information that was necessary  
for this analysis. These filtering criteria are very similar to those from Foote and  
<sub>122</sub> Miller (2013) with an additional constraint of being present in the body size  
data set from Payne et al. (2014). Epicontinental versus open-ocean assignments  
<sub>124</sub> for each fossil occurrence are partially based on those from Miller and Foote  
(2009), with additional occurrences assigned similarly (Miller and Foote,  
<sub>126</sub> personal communication).

Fossil occurrences were analyzed at the genus level which is common for  
<sub>128</sub> paleobiological, macroevolution, or macroecological studies of marine  
invertebrates (Alroy, 2010, Foote and Miller, 2013, Harnik et al., 2013, Kiessling  
<sub>130</sub> and Aberhan, 2007, Miller and Foote, 2009, Nürnberg and Aberhan, 2013, 2015,  
Payne and Finnegan, 2007, Simpson and Harnik, 2009, Vilhena et al., 2013).  
<sub>132</sub> While species diversity dynamics tend to be of much greater interest than those  
of higher taxa, the nature of the fossil record makes accurate and precise  
<sub>134</sub> taxonomic assignments at the species level for all occurrences extremely difficult  
if not impossible. Additionally, there is evidence of real differences in biological  
<sub>136</sub> patterns at the genus level versus the species level (Jablonski, 1987). As such,  
the choice to analyze genera as opposed to species was in order to assure a  
<sub>138</sub> minimum level of confidence and accuracy in the data analyzed here.

Genus duration was calculated as the number of geologic stages from first  
<sub>140</sub> appearance to last appearance, inclusive. Durations were based on geologic

stages as opposed to millions of years because of the inherently discrete nature  
142 of the fossil record; dates are not assigned to fossils themselves but instead  
fossils are known from a geological interval which represents some temporal  
144 range. Stages act as effectively irreducible globally consistent temporal intervals  
in which taxa occur.

146 Genera with a last occurrence in or after Changhsingian stage were right  
censored at the Changhsingian. Genera with a duration of only one stage were  
148 left censored (Appendix A). The covariates used to model genus duration were  
geographic range size ( $r$ ), environmental preference ( $v, v^2$ ), and body size ( $m$ ).

150 Geographic range was calculated using an occupancy approach. First, all  
occurrences were projected onto an equal-area cylindrical map projection. Each  
152 occurrence was then assigned to one of the cells from a  $70 \times 34$  regular raster  
grid placed on the map. Each grid cell represents approximately  $250,000 \text{ km}^2$ .  
154 The map projection and regular lattice were made using shape files from  
<http://www.naturalearthdata.com/> and the **raster** package for R (Hijmans,  
156 2015).

For each stage, the total number of occupied grid cells, or cells in which a fossil  
158 occurs, was calculated. Then, for each genus, the number of grid cells occupied  
by that genus was calculated. Dividing the genus occupancy by the total  
160 occupancy gives the relative occupancy of that genus. Mean relative genus  
occupancy was then calculated as the mean of the per stage relative occupancies  
162 of that genus.

Body size data was sourced directly from Payne et al. (2014). Because those  
164 measurements are presented without error, a measurement error model similar  
to the one for environmental affinity could not be implemented (Appendix A).

166 Prior to analysis, geographic range and body size were transformed and

standardized in order to improve interpretability of the results. Geographic  
168 range, which can only vary between 0 and 1, was logit transformed. Body size,  
which is defined for all positive real values, was natural log transformed. These  
170 covariates were then standardized by mean centering and dividing by two times  
their standard deviation following Gelman and Hill (2007).

## 172 2.2 Analytical approach

Hierarchical modelling is a statistical approach which explicitly takes into  
174 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  
176 units of study (e.g. genera) each belong to a single grouping (e.g. origination  
cohort). These groups are considered separate draws from a shared probability  
178 distribution (e.g. all cohorts, observed and unobserved). The group-level  
parameters are then estimated simultaneously as the other parameters of  
interest (e.g. covariate effects) (Gelman et al., 2013). The subsequent estimates  
180 are partially pooled together, where parameters from groups with large samples  
reflect our statistical uncertainty. Additionally, this partial pooling helps control  
182 for multiple comparisons and possibly spurious results as effects with little  
support are drawn towards the overall group mean.

184 This partial pooling is one of the greatest advantages of hierarchical modeling.  
By letting the groups “support” each other, parameter estimates then better  
186 reflect our statistical uncertainty. Additionally, this partial pooling helps control  
for multiple comparisons and possibly spurious results as effects with little  
188 support are drawn towards the overall group mean (Gelman et al., 2013,  
Gelman and Hill, 2007).

190 All covariate effects (regression coefficients), as well as the intercept term  
(baseline extinction risk), were allowed to vary by group (origination cohort).

<sup>192</sup> The covariance/correlation between covariate effects was also modeled. This  
<sup>193</sup> hierarchical structure allows inference for how covariates effects may change  
<sup>194</sup> with respect to each other while simultaneously estimating the effects  
themselves, propagating our uncertainty through all estimates.

<sup>196</sup> Additionally, instead of relying on point estimates of environmental affinity, I  
<sup>197</sup> treat environmental affinity as a continuous measure of the difference between  
<sup>198</sup> the taxon's environmental occurrence pattern and the background occurrence  
pattern (Appendix A).

### <sup>200</sup> 2.3 Survival model

<sup>202</sup> Genus durations were assumed to follow either an exponential or Weibull  
<sup>203</sup> distribution, both of which make different assumptions about how a taxon's  
<sup>204</sup> duration may effect its instantaneous extinction risk (Klein and Moeschberger,  
<sup>205</sup> 2003). The exponential distribution assumes that extinction risk is independent  
<sup>206</sup> of duration. In contrast, the Weibull distribution allows for age dependent  
<sup>207</sup> extinction via the shape parameter  $\alpha$ , though only as a monotonic function of  
<sup>208</sup> duration. Importantly, the Weibull distribution is equivalent to the exponential  
distribution when  $\alpha = 1$ .

<sup>210</sup> The following variables are here defined:  $y_i$  is the duration of genus  $i$  in geologic  
<sup>211</sup> stages,  $X$  is the matrix of covariates including a constant term,  $B_j$  is the vector  
<sup>212</sup> of regression coefficients for origination cohort  $j$ ,  $\Sigma$  is the covariance matrix of  
the regression coefficients,  $\tau$  is the vector of scales the standard deviations of  
<sup>214</sup> the between-cohort variation in regression coefficient estimates,  $\Omega$  is the  
correlation matrix of the regression coefficients, and  $\alpha_j$  is the shape parameter  
<sup>216</sup> for cohort  $j$  with  $a$  is the overall mean shape parameter and  $\pi$  is the variance  
between estimates of  $\alpha_j$ .

The exponential model is defined

$$y_i \sim \text{Exponential}(\lambda)$$

$$\lambda_i = \exp(\mathbf{X}_i B_{j[i]})$$

$$B \sim \text{MVN}(\vec{\mu}, \Sigma)$$

$$\Sigma = \text{Diag}(\vec{\tau}) \Omega \text{Diag}(\vec{\tau})$$

$$\mu_k \sim \begin{cases} \mathcal{N}(0, \psi_k \nu) & \text{if } k \neq r, \text{ or} \\ \mathcal{N}(-1, 1) & \text{if } k = r \end{cases} \quad (1)$$

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

$$\psi_k \sim C^+(1) \text{ if } k \neq r$$

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

$$\Omega \sim \text{LKJ}(2).$$

<sup>218</sup> Similarly, the Weibull model is defined

$$\begin{aligned}
y_i &\sim \text{Weibull}(\alpha_{j[i]}, \sigma) \\
\sigma_i &= \exp \left( \frac{-(\mathbf{X}_i B_{j[i]})}{\alpha_{j[i]}} \right) \\
B &\sim \text{MVN}(\vec{\mu}, \Sigma) \\
\Sigma &= \text{Diag}(\vec{\tau}) \Omega \text{Diag}(\vec{\tau}) \\
\log(\alpha) &\sim \mathcal{N}(a, \pi) \\
\mu_k &\sim \begin{cases} \mathcal{N}(0, \psi_k \nu) & \text{if } k \neq r, \text{ or} \\ \mathcal{N}(-1, 1) & \text{if } k = r \end{cases} \\
\tau_k &\sim \text{C}^+(1) \\
a &\sim \mathcal{N}(0, 1) \\
\pi &\sim \text{C}^+(1) \\
\psi_k &\sim \text{C}^+(1) \text{ if } k \neq r \\
\nu &\sim \text{C}^+(1) \\
\Omega &\sim \text{LKJ}(2).
\end{aligned} \tag{2}$$

The principal difference between this model and the previous (Eq. 1) is the  
<sup>220</sup> inclusion of the shape parameter  $\alpha$ . Note that  $\sigma$  is approximately equivalent to  
 $1/\lambda$ .

<sup>222</sup> For an explanation of how this model was developed, parameter explanations,  
and choice of priors, please see Appendix B. Note that these models (Eq. 1, 2)  
<sup>224</sup> do not include how the uncertainty in environmental affinity is included nor how  
censored observations are included. For an explanation of both of these aspects,  
<sup>226</sup> see Appendices A and C.

## 2.4 Parameter estimation

<sup>228</sup> The joint posterior was approximated using a Markov chain Monte Carlo  
<sup>229</sup> routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn  
<sup>230</sup> Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic  
<sup>231</sup> programming language Stan (Stan Development Team, 2014). The posterior  
<sup>232</sup> distribution was approximated from four parallel chains run for 10,000 draws  
<sup>233</sup> each, split half warm-up and half sampling and thinned to every 10th sample for  
<sup>234</sup> a total of 5000 posterior samples. Chain convergence was assessed via the scale  
<sup>235</sup> reduction factor  $\hat{R}$  where values close to 1 ( $\hat{R} < 1.1$ ) indicate approximate  
<sup>236</sup> convergence, meaning that the chains are approximately stationary and the  
samples are well mixed (Gelman et al., 2013).

## <sup>238</sup> 2.5 Model evaluation

<sup>240</sup> Models were evaluated using both posterior predictive checks and an estimate of  
out-of-sample predictive accuracy. The motivation behind posterior predictive  
checks as tools for determining model adequacy is that replicated data sets  
<sup>242</sup> using the fitted model should be similar to the original data (Gelman et al.,  
2013). Systematic differences between the simulations and observations indicate  
<sup>244</sup> weaknesses of the model fit. An example of a technique that is very similar  
would be inspecting the residuals from a linear regression.

<sup>246</sup> The strategy behind posterior predictive checks is to draw simulated values  
from the joint posterior predictive distribution,  $p(y^{rep}|y)$ , and then compare  
<sup>248</sup> those draws to the empirically observed values (Gelman et al., 2013). To  
accomplish this, for each replicate, a single value is drawn from the marginal  
<sup>250</sup> posterior distributions of each regression coefficient from the final model as well  
as estimates of  $\alpha_j$  for the Weibull model (Eq. 1, 2). Then, given the covariate

252 information  $\mathbf{X}$ , a new set of  $n$  genus durations are generated giving a single  
253 replicated data set  $y^{rep}$ . This is repeated 1000 times in order to provide a  
254 distribution of possible values that could have been observed given the model.

255 In order to compare the fitted model to the observed data, various graphical  
256 comparisons or test quantities need to be defined. The principal comparison  
257 used here is a comparison between non-parameteric approximation of the  
258 survival function  $S(t)$  as estimated from both the observed data and each of the  
259 replicated data sets. The purpose of this comparison is to determine if the  
260 model approximates the same survival/extinction pattern as the original data.

261 The exponential and Weibull models were compared for out-of-sample predictive  
262 accuracy using the widely-applicable information criterion (WAIC) (Watanabe,  
263 2010). Out-of-sample predictive accuracy is a measure of the expected fit of the  
264 model to new data. However, because the Weibull distribution reduces to the  
265 exponential distribution when  $\alpha = 1$ , my interest is not in choosing between  
266 these models. Instead, comparisons of WAIC values are useful for better  
267 understanding the effect of model complexity on out-of-sample predictive  
268 accuracy. The calculation of WAIC used here corresponds to the “WAIC 2”  
269 formulation recommended by Gelman et al. (2013). For an explanation of how  
270 WAIC is calculated, see Appendix D. Lower values of WAIC indicate greater  
expected out-of-sample predictive accuracy than higher values.

### 272 3 Results

273 As stated above, posterior approximations for both the exponential and Weibull  
274 models achieved approximate stationarity after 10,000 steps, as all parameter  
estimates have an  $\hat{R} < 1.1$ .

276 Comparisons of the survival functions estimated from 1000 posterior predictive  
277 data sets to the estimated survival function of the observed genera demonstrates  
278 that both the exponential and Weibull models approximately capture the  
279 observed pattern of extinction (Fig. 1). This is partially expected given that the  
280 unit of analysis is genus species duration as opposed species level which can  
281 alter the shape of  $S(t)$  (Foote, 2001, Raup, 1975, 1978, 1985). The major  
282 difference in fit between the two models is that the Weibull model has a slightly  
283 better fit for longer lived taxa than the exponential model.

284 Additionally, the Weibull model is expected to have slightly better  
285 out-of-sample predictive accuracy when compared to the exponential model  
286 (WAIC 4534 versus 4604). Because the difference in WAIC between these two  
287 models is large, only those results from the Weibull model will be discussed.

288 Estimates of the overall mean covariate effects  $\mu$  can be considered  
289 time-invariant generalizations for brachiopod survival during the Paleozoic (Fig.  
290 1). Consistent with prior expectations, geographic range size has a negative  
291 effect on extinction risk, where genera with large ranges having greater  
292 durations than genera with small ranges.

I infer that body size has no distinguishable effect on brachiopod taxonomic  
294 survival. The mean estimate of the effect of body size on extinction risk is  
295 negative, implying that increasing body size decreases extinction risk, however  
296 this estimate is within 2 standard deviations of 0 (mean  $\mu_m = -0.07$ , standard  
297 deviation 0.08; Fig. 1).

298 Interpretation of the effect of environmental preference  $v$  on duration is slightly  
299 more involved. Because a quadratic term is the equivalent of an interaction  
300 term, both  $\mu_v$  and  $\mu_{v^2}$  have to be interpreted together because it is illogical to  
301 change values of  $v$  without also changing values  $v^2$ . To determine the nature of

<sup>302</sup> the effect of  $v$  on duration I calculated the multiplicative effect of environmental preference on extinction risk.

<sup>304</sup> Given mean estimated extinction risk  $\tilde{\sigma}$ , we can define the extinction risk multiplier of an observation with environmental preference  $v_i$  as

$$f(v_i) = \exp\left(\frac{-(\mu_v v_i + \mu_{v^2} v^2)}{\exp(a)}\right). \quad (3)$$

<sup>306</sup> This function  $f(v_i)$  has a y-intercept of  $\exp(0)$  or 1 because it does not have a non-zero intercept term. Equation 3 can be either concave up or down. A <sup>308</sup> concave down  $f(v_i)$  may indicate that genera of intermediate environmental preference have greater durations than either extreme, and *vice versa* for <sup>310</sup> concave up function.

The expected effect of environmental preference as a multiplier of expected <sup>312</sup> extinction risk can then be visualized (Fig. 2). This figure depicts 1000 posterior predictive estimates of Eq. 3 across all possible values of  $v$ . The number <sup>314</sup> indicates the posterior probability that the function is concave down, with generalists having lower extinction risk/greater duration than either type of <sup>316</sup> specialist. Note that the inflection point of Fig. 2 is approximately  $x = 0$ , something that is expected given the estimate of  $\mu_v$  (Fig. 1).

<sup>318</sup> The matrix  $\Sigma$  describing the covariance between the different coefficients describes how these coefficients might vary together across the origination <sup>320</sup> cohorts. Similar to how this was modeled (Eq. 1, 2), for interpretation purposes  $\Sigma$  is decomposed into a vector of standard deviations  $\vec{\tau}$  and a correlation matrix <sup>322</sup>  $\Omega$ .

The estimates of the standard deviation of between-cohort coefficient estimates <sup>324</sup>  $\tau$  indicate that some effects can vary greatly between-cohorts (Fig. 1).

Coefficients with greater values of  $\tau$  have greater between-cohort variation. The

326 covariate effects with the greatest between origination cohort variation are  $\beta_r$ ,  
327  $\beta_v$ , and  $\beta_{v^2}$ . Estimates of  $\beta_m$  have negligible between cohort variation, as there  
328 is less between cohort variation than the between cohort variation in baseline  
extinction risk  $\beta_0$ . However the amount of between cohort variation in estimates  
329 of  $\beta_{v^2}$  means that it is possible for the function describing the effect of  
environmental affinity to be upward facing for some cohorts (Eq. 3), which  
330 corresponds to environmental generalists being shorter lived than specialists in  
that cohort.

334 The correlation terms of  $\Omega$  (Fig. 3a) describe the relationship between the  
coefficients and how their estimates may vary together across cohorts. The  
336 correlations between the intercept term  $\beta_0$  and the effects of the taxon traits are  
of particular interest for evaluating the Jablonski (1986) scenario (Fig. 3a first  
338 column/last row). Keep in mind that when  $\beta_0$  is low, extinction risk is low; and  
conversely, when  $\beta_0$  is high, then extinction risk is high.

340 Marginal posterior probabilities of the correlations between the level of baseline  
extinction risk  $\beta_0$  and the effects of the taxon traits indicate that the correlation  
342 between expected extinction risk and both geographic range  $\beta_r$  and  $\beta_{v^2}$  are of  
particular note (Fig. 3b).

344 There is only an approximately 85% probability that  $\beta_0$  and  $\beta_r$  are negatively  
correlated (Fig. 3b), meaning that as extinction risk increases, the  
346 effect/importance of geographic range on genus duration may increases. This  
means that increases in baseline extinction rate are weakly correlated with an  
348 increased importance of geographic range size. There is a 93% probability that  
 $\beta_0$  and  $\beta_{v^2}$  are negatively correlated (Fig. 3b), meaning that as extinction risk  
350 decreases, the peakedness of  $f(v_i)$  may increases and the relationship would  
then tend towards concave down. Additionally, there is a 94% probability that  
352 values of  $\beta_r$  and  $\beta_{v^2}$  are positively correlated (Mean correlation 0.45, standard

deviation 0.26).

354 While the overall group level estimates are of particular importance when  
defining time-invariant differences in extinction risk, it is also important and  
356 useful to analyze the individual level parameter estimates in order to better  
understand how parameters actually vary across cohorts.

358 In comparison to the overall mean extinction risk  $\mu_0$ , cohort level estimates  $\beta_0$   
show some amount of variation through time as expected by estimates of  $\tau_0$   
360 (Fig. 4). A similar, if slightly greater, amount of variation is also observable in  
cohort estimates of the effect of geographic range  $\beta_r$  (Fig. 4). Again, smaller  
362 values of  $\beta_0$  correspond to lower expected extinction risk. Similarly, smaller  
values of  $\beta_r$  correspond to greater decrease in extinction risk with increasing  
364 geographic range

How the effect of environmental affinity varies between cohorts can be observed  
366 by using the cohort specific coefficients estimates. Following the same procedure  
used earlier (Fig. 1), but substituting cohort specific estimates of  $\beta_v$  and  $\beta_{v^2}$  for  
368  $\mu_v$  and  $\mu_{v^2}$ , the cohort specific effect of environmental preference as a multiplier  
of mean extinction risk can be calculated. This was done only for the Weibull  
370 model, though the observed pattern should be similar for the exponential model.

As expected based on the estimates of  $\tau_v$  and  $\tau_{v^2}$ , there is greater variation in  
372 the peakedness of  $f(v_i)$  than there is variation between convave up and down  
functions (Fig. 5). 12 of the 33 cohorts have less than 50% posterior probability  
374 that generalists are potentially expected to be shorter lived than specialists,  
though two of those cases have approximately a 50% probability of being either  
376 concave up or down. This is congruent with the 0.74 posterior probability that  
 $\mu_{v^2}$  is positive/ $f(v_i)$  is concave down.

378 Additionally, for some cohorts there is a quite striking pattern where the effect

of environmental preference  $v$  has a nearly-linear relationship (Fig. 5). These are  
380 primarily scenarios where one of the end member preferences is expected to  
have a greater duration than either intermediate or the opposite end member  
382 preference. Whatever curvature is present in these nearly-linear cases is due to  
the definition of  $f(v)$  as it is not defined for non-negative values of  $\sigma$  (Eq. 3).  
384 For all stages between the Emsian through the Viséan, inclusive, intermediate  
preferences are of intermediate extinction risk when compared with  
386 epicontinental specialists (lowest risk) or open-ocean specialists (highest risk).  
This time period represents most of the Devonian through the early  
388 Carboniferous.

Interestingly, the estimates of  $\alpha_j$  have low between cohort variation with most  
390 cohorts having approximately the same estimate as the overall estimate of  $\alpha$   
(Fig. 4). Also, for the majority of the time analyzed there is evidence for  
392 accelerating extinction with taxon age. If this is due to biological causes  
(Rosindell et al., 2015, Wagner and Estabrook, 2014) or an artifact of  
394 preservation/the minimum observable duration of a genus (Sepkoski, 1975) is  
unknown. While it is expected that a non-exponential model will better fit  
396 genus-level data, this would be reflected as decelerating extinction risk with  
taxon age as opposed to accelerating (Foote, 2001, Raup, 1975, 1978, 1985).  
398 This uncertainty remains an open issue for future analysis.

## 4 Discussion

400 My results demonstrate that both the effects of geographic range and the  
peakedness/concavity of environmental preference are both negatively  
402 correlated with baseline extinction risk, meaning that as baseline extinction risk  
increases the effect size of geographic range increases but the curvature of the

404 effect of environmental preference decreases (Fig. 3b).

The above results provide slightly more support for the hypothesis that as  
406 extinction intensity increases, the effects of biological traits that aren't  
geographic range tend decrease in magnitude. However, the interaction between  
408 geographic range and origination cohort, while weak, indicates that neither of  
the two proposed end-member macroevolutionary mechanisms is strictly right.

410 The evidence for a correlation between changes in the effect of geographic range  
and the curvature of the effect of environmental preference may help explain  
412 this result. I hypothesize that because taxa with large geographic ranges  
encompass more possible environments, when extinction intensity is high there  
414 is little coherent difference in the environmental preference among the surviving  
taxa. The intensity decreases the selectivity such that the effect of  
416 environmental preference is effectively washed out by the strength of the effect  
of geographic range.

418 For the approximately 233 My period analyzed there is an approximate 75%  
posterior probability that brachiopod genera with intermediate environmental  
420 preferences are expected to have a lower extinction risk than either end  
members. However, the over all curvature of  $f(v_i)$  is not very peaked meaning  
422 that when averaged over the entire Phanerozoic this relationship may not lead  
to large differences in extinction risk (Fig. 2). Note that the duration of the  
424 period analyzed is approximately four times then length of the Cenozoic (e.g.  
time since the extinction of the non-avian dinosaurs). This result gives weak  
426 support for the universality of the hypothesis that environmental generalists  
have greater survival than environmental specialists (Liow, 2004, 2007,  
428 Nürnberg and Aberhan, 2013, 2015, Simpson, 1944).

The posterior variance in the estimate of overall  $f(v_i)$  reflects the large between

430 cohort variance in cohort specific estimates of  $f(v_i)$  (Fig. 5). Given that there is  
only a 75% posterior probability that the expected overall estimate of  $f(v_i)$  is  
432 concave down, it is not surprising that there are some stages where the  
estimated relationship is in fact the reverse of the prior expectation.

434 Additionally, some of those same stages where  $f(v_i)$  does not resemble the prior  
expectation of a concave down nonlinear relation are instead highly skewed  
436 and effectively linear (Fig. 5). These results demonstrate that, while the  
group-level estimate may only weakly support one hypothesis, the cohort-level  
438 estimates may exhibit very different characteristics. These results are also  
consistent with aspects of Miller and Foote (2009) who found that the effect of  
440 environmental preference on extinction risk was quite variable and without  
obvious patterning during times of background extinction.

442 There are two mass extinction events that are captured within the time frame  
considered here: the Ordovician-Silurian and the Frasnian-Famennian. The  
444 cohorts bracketing these events are worth considering in more detail.

The proposed mechanism for the end Ordovician mass extinction is a decrease  
446 in sea level and the draining of epicontinental seas due to protracted glaciation  
(Johnson, 1974, Sheehan, 2001). My results are broadly consistent with this  
448 scenario with both epicontinental and open-ocean specialists having a much  
lower expected duration than intermediate taxa (Fig. 5). All of the stages  
450 between the Darriwillian and the Llandovery, except the Hirnantian, have a  
high probability (90+%) that  $f(v)$  is concave down. The pattern for the  
452 Darriwillian, which proceeds the supposed start of Ordovician glacial activity,  
demonstrates that taxa tend to favor open-ocean environments are expected to  
454 have a greater duration than either intermediate or epicontinental specialists, in  
decreasing order.

456 For nearly the entire Devonian estimates of  $f(v)$  indicate that one of the

environmental end members is favored over the other end member of  
458 intermediate preference (Fig. 5). For almost the entire Givetian through  
the end of the Devonian and into the Viséan, I find that epicontinental favoring  
460 taxa are expected to have a greater duration than either intermediate or  
open-ocean specialists. Additionally, for nearly the entire Devonian and through  
462 to the Visean, the cohort-specific estimates of  $f(v)$  are concave-up. This is the  
opposite pattern than what is expected (Fig. 2). This result, however, seems to  
464 reflect the intensity of the seemingly nearly-linear difference in expected  
duration across the range of  $v$  as opposed to an inversion of the weakly expected  
466 curvilinear pattern.

The use of genera as the unit of the study and how to exactly interpret the  
468 effects of the biological traits is a remaining concern. For example, if any of the  
traits analyzed here are associated with increases in speciation rates, this might  
470 increase the duration of genera through self-renewal (Raup, 1991, 1994), which  
would be an example of the difference in biological pattern between species and  
472 genera (Jablonski, 1987, 2007, 2008). This could lead to a trait appearing to  
decrease generic level extinction risk by increasing species level origination rate  
474 instead of decreasing species level extinction risk. However, given the nature of  
the brachiopod fossil record and the difficulty of identifying individual  
476 specimens to the species level, there is no simple solution to decreasing this  
uncertainty in the interpretations of how the biological traits studied here  
478 actually affect extinction risk.

This model could be improved through either increasing the number of analyzed  
480 taxon traits, expanding the hierarchical structure of the model to include other  
major taxonomic groups of interest, and the inclusion of explicit phylogenetic  
482 relationships between the taxa in the model as an additional hierarchical effect.

An example taxon trait that may be of particular interest is the affixing

<sup>484</sup> strategy or method of interaction with the substrate of the taxon. This trait has  
been found to be related to brachiopod survival (Alexander, 1977) so its  
<sup>486</sup> inclusion may be of particular interest.

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

<sup>494</sup> Taxon traits like environmental preference or geographic range (Hunt et al.,  
2005, Jablonski, 1987) are most likely heritable, at least phylogenetically  
<sup>496</sup> (Housworth et al., 2004, Lynch, 1991). Without phylogenetic context, this  
analysis assumes that differences in extinction risk between taxa are  
<sup>498</sup> independent of those taxa's shared evolutionary history (Felsenstein, 1985). In  
contrast, the origination cohorts only capture shared temporal context. The  
<sup>500</sup> inclusion of phylogenetic context as an addition individual level hierarchical  
structure independent of origination cohort would allow for determining how  
<sup>502</sup> much of the observed variability is due to shared evolutionary history versus  
actual differences associated with these taxonomic traits.

<sup>504</sup> In summary, patterns of Paleozoic brachiopod survival were analyzed using a  
fully Bayesian hierarchical survival modelling approach while also eschewing the  
<sup>506</sup> traditional separation between background and mass extinction. I modeled both  
the overall mean effect of biological covariates on extinction risk while also  
<sup>508</sup> modeling the correlation between cohort-specific estimates of covariate effects. I  
find that as baseline extinction risk increases, the form of the selectivity of  
<sup>510</sup> extinction changes such that during periods of low extinction the effect

environmental preference has a more nonlinear relationship than during times of  
512 high extinction intensity. In particular, the correlation between the effect of  
geographic range and the curvature of the effect of environmental preference on  
514 taxon survival supports the idea that during periods of high extinction intensity  
the size of effect of geographic range effectively washes out the effects of other  
516 biological traits. Finally, I find very weak for “survival of the unspecialized”  
(Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944) as a  
518 general characterization of the effect of environmental preference on extinction  
risk (Fig. 2), though there is heterogeneity between origination cohorts with  
520 most periods of time conforming to this hypothesis (Fig. 5).

## Acknowledgements

522 I would like to thank K. Angielczyk, M. Foote, P. D. Polly, and R. Ree for  
helpful and Marites Marites Villarosa Garcia for draft comments. Additionally,  
524 thank you A. Miller for the epicontinental versus open-ocean assignments. This  
entire study would not have been possible without the Herculean effort of  
526 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,  
528 A. Hendy, S. Holland, L. Ivany, W. Kiessling, B. Kröger, A. McGowan, T.  
Olszewski, P. Novack-Gottshall, M. Patzkowsky, M. Uhen, L. Villier, and P.  
530 Wager. This work was supported by a NASA Exobiology grant (NNX10AQ446)  
to A. Miller and M. Foote. I declare no conflicts of interest. This is Paleobiology  
532 Database publication XXX.

## References

- 534 Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation  
to the mode of stabilization on the substrate. *Palaeogeography,*  
536 *Palaeoclimatology, Palaeoecology* 21:209–226.
- Alroy, J., 2010. The Shifting Balance of Diversity Among Major Marine Animal  
538 Groups. *Science* 329:1191–1194.
- Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of  
540 filter morphology on evolutionary rates. *Paleobiology* 19:304–321.
- Cooper, W. S., 1984. Expected time to extinction and the concept of  
542 fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- Felsenstein, J., 1985. Phylogenies and the comparative method. *American  
544 Naturalist* 125:1–15.
- Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites.  
546 *Paleobiology* 14:258–271.
- , 2000. Origination and extinction components of taxonomic diversity:  
548 Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- , 2001. Evolutionary Rates and the Age Distributions of Living and  
550 Extinct Taxa. Pp. 245–294, *in* J. B. Jackson, S. Lidgard, and F. K.  
McKinney, eds. *Evolutionary Patterns: Growth, Form, and Tempo in the  
552 Fossil Record*. University of Chicago Press, Chicago.
- Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine  
554 animal genera. *Paleobiology* 39:171–192.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B.

- 556 Rubin, 2013. Bayesian data analysis. 3 ed. Chapman and Hall, Boca Raton,  
FL.
- 558 Gelman, A. and J. Hill, 2007. Data Analysis using Regression and  
Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- 560 Harnik, P. G., C. Simpson, and J. L. Payne, 2013. Long-term differences in  
extinction risk among the seven forms of rarity. Proceedings of the Royal  
562 Society B: Biological Sciences 282.
- Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL  
564 <http://CRAN.R-project.org/package=raster>. R package version 2.3-24.
- 566 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.
- 568 Housworth, E. A., P. Martins, and M. Lynch, 2004. The Phylogenetic Mixed  
Model. The American Naturalist 163:84–96.
- 570 Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a  
comment on "On the heritability of geographic range sizes". American  
572 Naturalist 166:129–135.
- Jablonski, D., 1986. Background and mass extinctions: the alternation of  
574 macroevolutionary regimes. Science 231:129–133.
- , 1987. Heritability at the species level: analysis of geographic ranges of  
576 cretaceous mollusks. Science 238:360–363.
- , 2007. Scale and hierarchy in macroevolution. Palaeontology 50:87–109.
- 578 ———, 2008. Species Selection: Theory and Data. Annual Review of Ecology,  
Evolution, and Systematics 39:501–524.

- 580 Jablonski, D. and K. Roy, 2003. Geographical range and speciation in fossil and  
living molluscs. *Proceedings. Biological sciences / The Royal Society*  
582 270:401–6.
- Johnson, J. G., 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- 584 Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine  
benthic biodiversity dynamics through Triassic–Jurassic time. *Paleobiology*  
586 33:414–434.
- Klein, J. P. and M. L. Moeschberger, 2003. *Survival Analysis: Techniques for  
588 Censored and Truncated Data*. 2nd ed. Springer, New York.
- Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively  
590 unspecialized" using fossil crinoids. *The American Naturalist* 164:431–43.
- , 2007. Does versatility as measured by geographic range, bathymetric  
592 range and morphological variability contribute to taxon longevity? *Global  
Ecology and Biogeography* 16:117–128.
- 594 Lynch, M., 1991. Methods for the analysis of comparative data in evolutionary  
biology. *Evolution* 45:1065–1080.
- 596 Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings:  
the kinetics of mass extinction and origination. *Science* 326:1106–9.
- 598 Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range  
predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology*  
600 39:360–372.
- , 2015. Interdependence of specialization and biodiversity in Phanerozoic  
602 marine invertebrates. *Nature communications* 6:6602.

- Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental  
604 than evolvability. *PLoS one* 7:e38025.
- Payne, J. L. and S. Finnegan, 2007. The effect of geographic range on  
606 extinction risk during background and mass extinction. *Proceedings of the  
National Academy of Sciences* 104:10506–11.
- 608 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  
610 million years. *Proceedings of the Royal Society B* 281:20133122.
- 612 Peters, S. E., 2008. Environmental determinants of extinction selectivity in the  
fossil record. *Nature* 454:626–9.
- Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law.  
614 *Paleobiology* 1:82–96.
- , 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- 616 ———, 1985. Mathematical models of cladogenesis. *Paleobiology* 11:42–52.
- , 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- 618 ———, 1994. The role of extinction in evolution. *Proceedings of the National  
Academy of Sciences* 91:6758–6763.
- 620 Rosindell, J., L. J. Harmon, and R. S. Etienne, 2015. Unifying ecology and  
macroevolution with individual-based theory. *Ecology letters* 18:472–82.
- 622 Sepkoski, J. J., 1975. Stratigraphic biases in the analysis of taxonomic  
survivorship. *Paleobiology* 1:343–355.
- 624 Sepkoski Jr., J. J., 1981. A factor analytic description of the Phanerozoic  
marine fossil record. *Paleobiology* 7:36–53.

- 626 Sheehan, P., 2001. The late Ordovician mass extinction. *Annual Review of Earth and Planetary Sciences* 29:331–364.
- 628 Simpson, C., 2006. Levels of selection and large-scale morphological trends. Ph.D. thesis, University of Chicago.
- 630 Simpson, C. and P. G. Harnik, 2009. Assessing the role of abundance in marine bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.
- 632 Simpson, G. G., 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.
- 634 ———, 1953. *The Major Features of Evolution*. Columbia University Press, New York.
- 636 Stan Development Team, 2014. Stan: A c++ library for probability and sampling, version 2.5.0. URL <http://mc-stan.org/>.
- 638 Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- , 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- 640 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
- 642 latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Scientific Reports* 3:1790.
- 644 Wagner, P. J. and G. F. Estabrook, 2014. Trait-based diversification shifts reflect differential extinction among fossil taxa. *Proceedings of the National*
- 646 *Academy of Sciences* 111:16419–16424.
- Wang, S. C., 2003. On the continuity of background and mass extinction. *Paleobiology* 29:455–467.
- 648 Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and

- <sup>650</sup> Widely Applicable Information Criterion in Singular Learning Theory.  
Journal of Machine Learning Research 11:3571–3594.

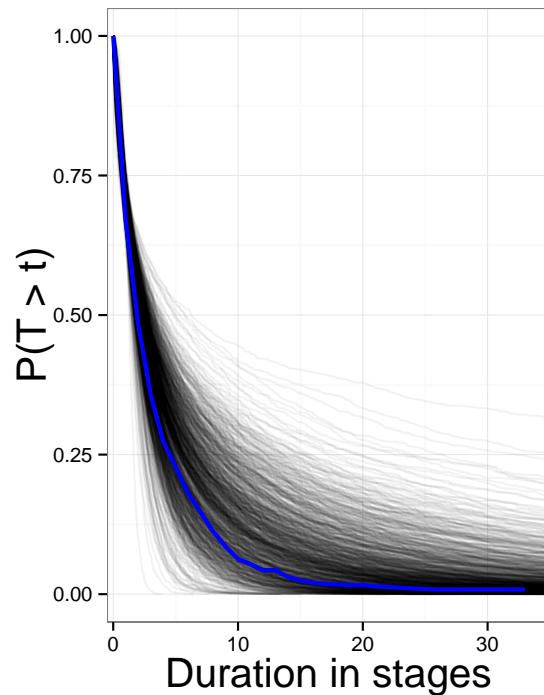


Figure 1: Comparison of empirical estimates of  $S(t)$  versus estimates from 1000 posterior predictive data sets.  $S(t)$  corresponds to  $P(T > t)$  as it is the probability that a given genus observed at age  $t$  will continue to live. This is equivalent to the probability that  $t$  is less than the genus' ultimate duration  $T$ .

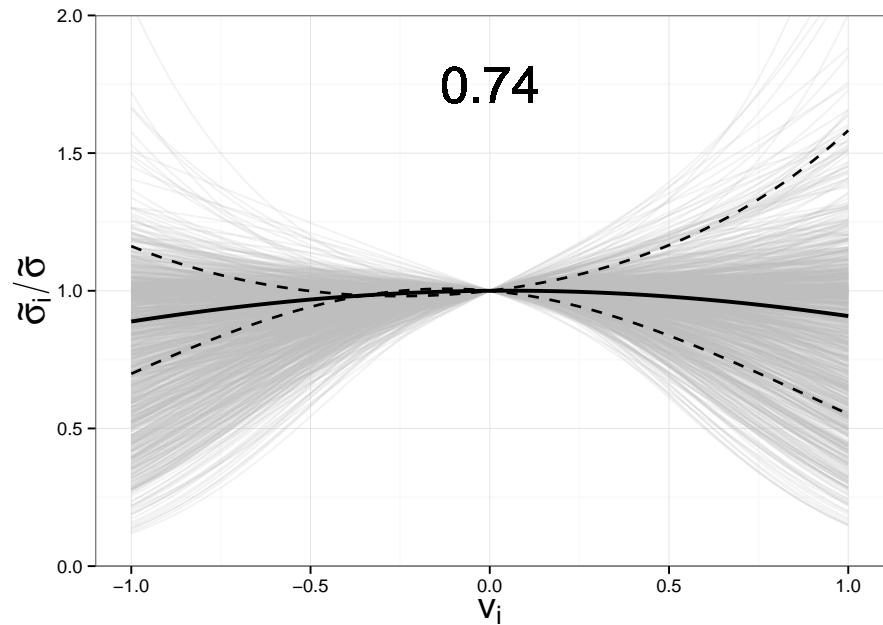


Figure 2: The overall expected relationship  $f(v_i)$  between environmental affinity  $v_i$  and a multiplier of extinction risk (Eq. 3). Each grey line corresponds to a single draw from the posterior predictive distribution, the solid black line corresponds to the median of the posterior predictive distribution, and the dashed black lines correspond to the median relationship plus or minus one standard deviation. The overall shape of  $f(v_i)$  is concave down with an optimum of close 0, which corresponds to affinity approximately equal to the expectation based on background environmental occurrence rates.

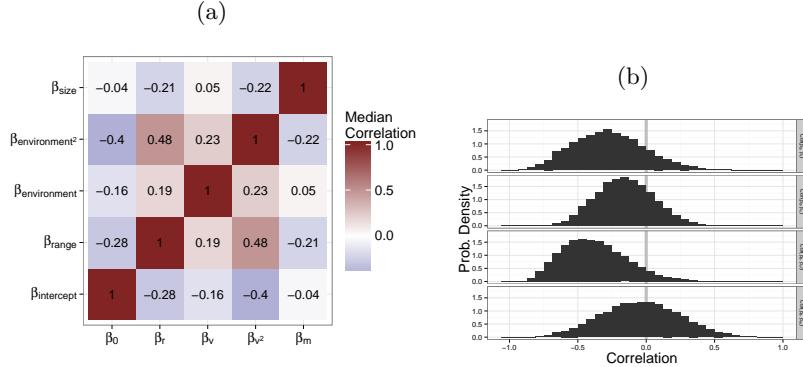


Figure 3: **A:** Heatmap for the median estimates of the terms of the correlation matrix  $\Omega$  between cohort-level covariate effects. Both the exponential (left) and Weibull (right) models are presented. The off-diagonal terms are the correlation between the estimates of the cohort-level estimates of the effects of covariates, along with intercept/baseline extinction risk. **B:** Marginal posterior distributions of the correlations between intercept terms/baseline extinction risk and the effects of each of the covariates. These are presented for both the exponential (left) and Weibull (right) models.

parameter	mean	standard deviation
$\mu_i$	-1.52	0.16
$\mu_r$	-1.39	0.13
$\mu_v$	-0.04	0.16
$\mu_{v^2}$	0.30	0.45
$\mu_m$	-0.07	0.08
$\tau_i$	0.77	0.14
$\tau_r$	0.40	0.13
$\tau_v$	1.05	0.23
$\tau_{v^2}$	1.87	0.64
$\tau_m$	0.24	0.13

Table 1: Group-level estimates of the intercept terms the effects of biological traits on brachiopod generic survival from equations 1 and 2, presented as means and standard deviations.  $\mu$  values are the location parameters of the effects, while  $\tau$  values are the scale terms describing the variation between cohorts. The subscripts correspond to the following:  $i$  intercept,  $r$  geographic range,  $v$  environmental affinity,  $v^2$  environmental affinity squared,  $m$  body size.

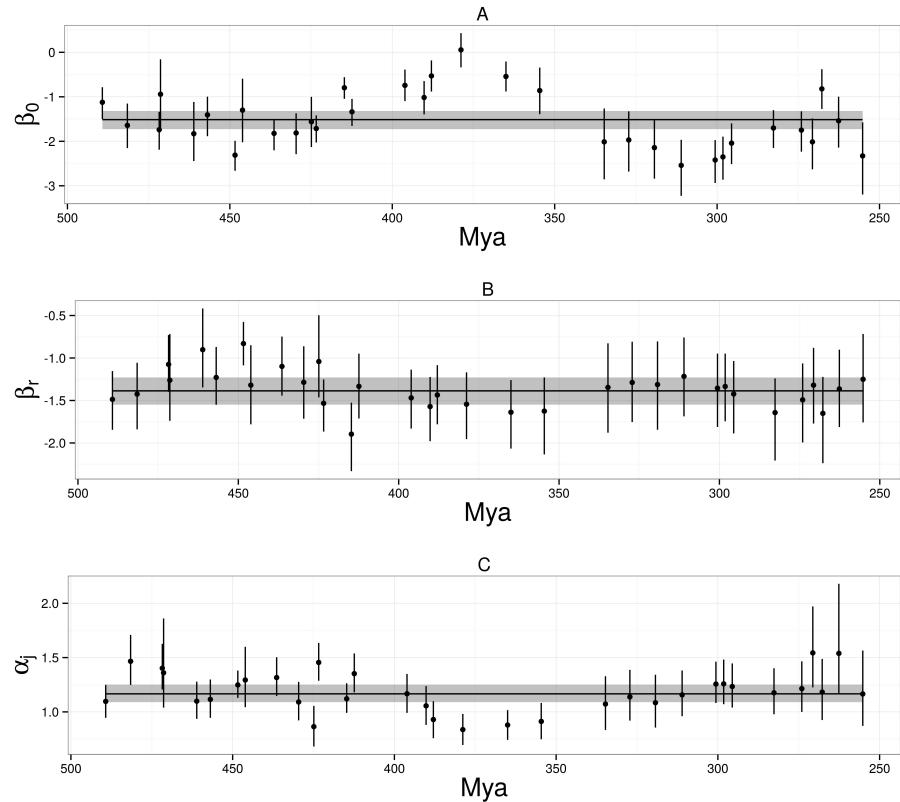


Figure 4: A) Comparison of cohort-specific estimates of  $\beta_0$  presented along with the estimate for the overall baseline extinction risk. B) Comparison of cohort-specific estimates of the effect of geographic range on extinction risk  $\beta_r$  presented along with the estimate for the overall effect of geographic range. C) Comparison of cohort-specific estimates of the Weibull shape parameter  $\alpha$ . Values greater than 1 correspond to accelerating extinction with age, and those below 1 to decelerating extinction with age. Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. The horizontal line is the median estimate of the overall baseline extinction risk along with 80% credible intervals.

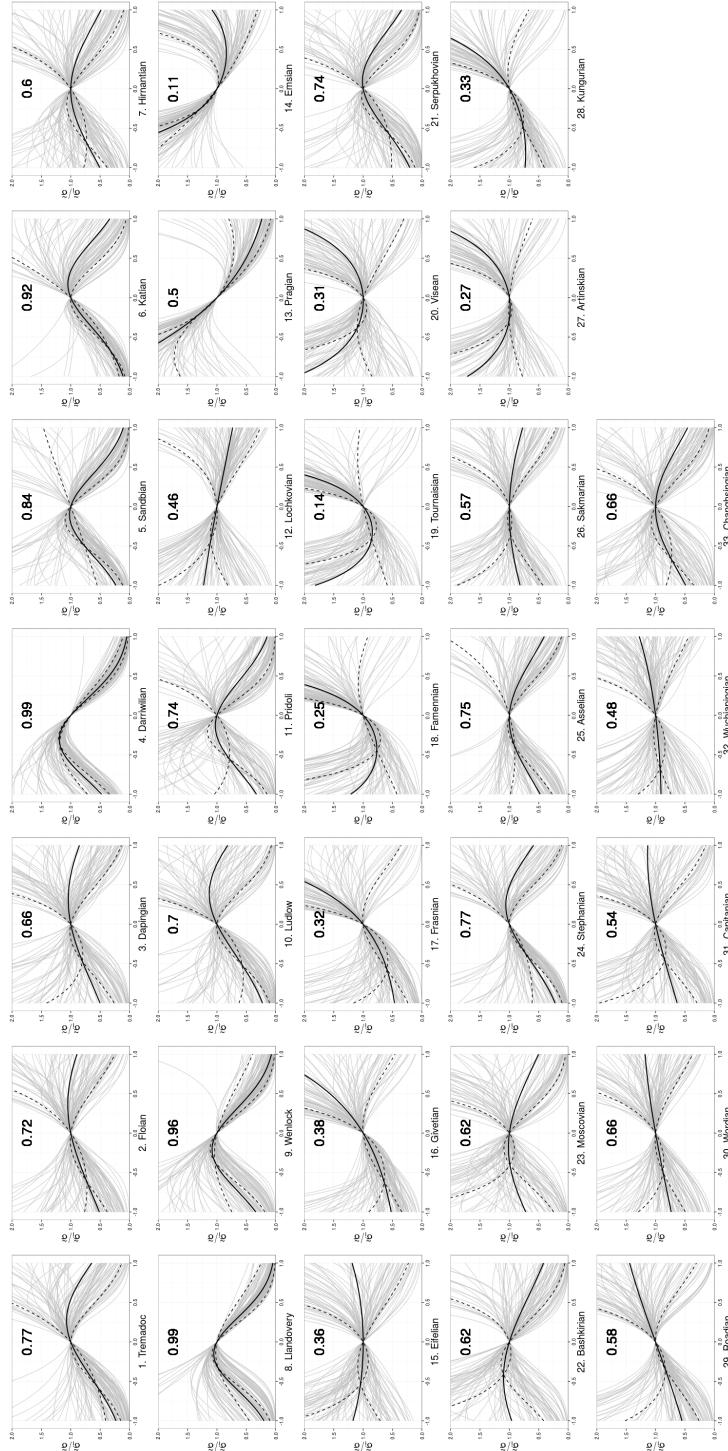


Figure 5: Comparison of the cohort-specific estimates of  $f(v_i)$  (Eq. 3) for the 33 analyzed origination cohorts. The stage of origination is labeled on the x-axis of each panel. The oldest stage is in the upper left, while the youngest is in the lower left. The number in each panel corresponds to the posterior probability that  $f(v_i)$  is concave down. Those that are highlighted in red have less than 51% posterior predictive probability that  $f(v_i)$  is concave down. Each grey line corresponds to a single draw from the posterior predictive distribution, the solid black line corresponds to the median of the posterior predictive distribution, and the dashed black lines correspond to the median relationship plus or minus one standard deviation. Note that all estimates must pass through  $y = 1$  when  $x = 0$  (Eq. 3).