

# 1 Manuscript

2 Justin Pomeranz<sup>1,\*</sup> James R. Junker<sup>2,3</sup> Jeff S. Wesner<sup>4</sup>

3 04 September, 2022

## 4 Contents

5 <b>Abstract</b>	2
6 <b>Introduction</b>	2
7 <b>Methods</b>	3
8 Data Simulation . . . . .	3
9 Finer resolution of $\lambda$ values . . . . .	3
10 Estimation of Size spectra parameter $\lambda$ . . . . .	4
11 Estimation of relationship across gradient, $\beta_1$ . . . . .	4
12 Empirical Data . . . . .	4
13 <b>Results</b>	4
14 Relationship across the hypothetical environmental gradient . . . . .	4
15 No relationship . . . . .	10
16 Small variation in lambda . . . . .	10
17 Empirical data . . . . .	10
18 Simulating shallower lambdas . . . . .	14
19 <b>Discussion</b>	16
20 Concluding Remarks . . . . .	16
21 <b>References</b>	18
22 <b>Supplementary material</b>	19
23 Range of body sizes, $M$ . . . . .	19
24 Large environmental gradient . . . . .	21
25 Varying number of sites . . . . .	24
26 Sample size . . . . .	30
27 $\lambda$ and relationship estimates . . . . .	34

28 <sup>1</sup> Colorado Mesa University, Grand Junction, CO, USA

29 <sup>2</sup> Great Lakes Research Center, Michigan Technological University, Houghton, MI USA

30 <sup>3</sup> Louisiana Universities Marine Consortium, Chauvin, LA USA

31 <sup>4</sup> Dept. of Biology, University of South Dakota, Vermillion, SD, USA

32 \* Correspondence: Justin Pomeranz <jfpomeranz@gmail.com>

33 **NOTE** I need to integrate reference manager with R markdown. Any favorites or recommendations?? I  
34 generally use Zotero (free), but think I have access to Mendeley through CMU...

35 **NOTE from Jim** I use Zotero, mainly for free and open-source reasons, but I also can have a stan-  
36 dalone program that doesn't require internet access, can backup my reference library on the cloud and sync  
37 among multiple computers, can spit out bibtex to connect to markdown, and can have group citation lists  
38 if necessary. I use it with a betterBibtex add on that allows for modifying bibtex keys for calling refs in  
39 Rmarkdown.

40       **NOTE** I re-ran all of these after setting the seed for reproducibility. Figures should be up to date, but  
41       the numbers may change slightly i.e., 94.8% might become 95.1%, mean  $\pm$  SD values will be different.

## 42       Abstract

- 43       1. Size spectra represent a fundamental attribute of community organization and are increasingly being  
44       used in assessments of fresh waters.
- 45       2. Many methods have been proposed for constructing size spectra relationships, but recent work has  
46       shown that varied methods return biased estimates of relationship parameters, and different methods  
47       in fact are not estimating the same parameter. Despite this variability in estimates, it is unclear if  
48       the relative change across environmental gradients is consistent across methodologies. Here, we simu-  
49       late data sets across an hypothetical environmental gradient and estimate the size spectra parameter  
50       (slope or exponent, depending on method) at each site, and, importantly, estimate the relationship of  
51       parameters across the hypothetical gradient. We also use two previously published body size datasets  
52       across an anthropogenic stress gradient and an environmental temperature gradient and assess how  
53       the conclusions of these studies would vary based on methods used.
- 54       3. We find that maximum likelihood methods always perform as well or better than common binning  
55       methods. Additionally, the variance in estimates using MLE methods is markedly reduced when  
56       compared to binning methods.
- 57       4. The uncertainty and variation in estimates when using binning methods is often greater than or equal  
58       to the variation previously published in experimental and observational studies, bringing into question  
59       the effect size of previously published results. However, re-analysis of two previously published datasets  
60       does not markedly alter the conclusions reached. Further study is needed to identify how and when  
61       these estimates may be biased, and when the general pattern of results can be considered to be “true”.

## 62       Introduction

63       Body size distributions are a fundamental characteristic of communities. The remarkable consistency of  
64       these relationships across spatiotemporal scales and habitat types has led them to be recommended as a “uni-  
65       versal” indicator of ecological status (Petchey and Belgrano 2010)(Petchey and Belgrano 2010). Size metrics  
66       have commonly been used in marine systems, and are increasingly being applied to assess the condition of  
67       freshwater ecosystems (Martinez et al. 2016, Pomeranz 2018)(Pomeranz et al. 2019, **martinez2016?**).

68       Individual size distributions (ISD, also referred to as abundance size spectra) are commonly used. Gen-  
69       erally, there is a negative relationship between body size ( $M$ ) and abundance ( $N$ ). Theoretical and empirical  
70       data support this relationship being described as a simple power law with exponent  $\lambda$  in the form of  $N \sim M^\lambda$ .  
71       Commonly,  $N$  is the count of body sizes grouped into bins, and  $\lambda$  is estimated from OLS regressions of log  
72       transformed data of  $N$  and the mid point of the body size bin  $M_{bin}$ :  $\log(N_{count}) = \lambda \log(M_{bin})$ . Myriad  
73       binning methods have been proposed, including linear and logarithmic bin widths. Likewise, methods can  
74       rely on the absolute count in the bins as well as normalization techniques where the count is divided by the  
75       bin width (especially common with logarithmic binning). Alternatively,  $\lambda$  can be estimated directly using  
76       maximum likelihood techniques.

77       Previous work has shown that the estimates of  $\lambda$  differ between MLE and OLS techniques. OLS methods  
78       are particularly sensitive to decisions made in the binning process. Simulation studies have shown that MLE  
79       offers consistently more accurate estimates of  $\lambda$  (White et al. 2007, Edwards et al. 2017)(White et al. 2007,  
80       Edwards et al. 2017), and reanalysis of empirical datasets also indicates that the conclusions are dependent  
81       on the methodology used (White et al. 2007, Edwards et al. 2020)(White et al. 2007, Edwards et al. 2020).  
82       However, recent empirical analysis of stream macroinvertebrate communities across the NAtional Ecological  
83       Observatory Network (NEON, USA) showed that while the estimates of  $\lambda$  varied, the relative change across  
84       the environmental gradient was consistent regardless of method used (Pomeranz et al. 2022)(Pomeranz et  
85       al. 2022). While there is a growing consensus that MLE methods offer more reliable estimates of  $\lambda$ , and  
86       binning methods result in biased estimates, it remains unclear if these biases are consistent and systematic or  
87       stochastic, and whether or not the relative change in ISD parameters is consistent across space and time. In  
88       other words, if the data within a study are all treated the same, does a relative change of OLS slope parameters

89 of 0.1 coincide with a relative change of MLE estimates of 0.1? In order to answer this question, we simulate  
90 body size observations from bounded power law distributions with varied  $\lambda$  exponents across a hypothetical  
91 environmental gradient and compare the results obtained using three different methodologies common in  
92 the literature. In addition, we re-analyze two previously published datasets of stream community body sizes  
93 across a stress and environmental gradient. We find that the MLE method more accurately estimates the  
94 site-specific  $\lambda$  exponents as well as the relative change in exponents across the hypothetical gradient, as  
95 well as having smaller variation. The logarithmic binning methods generally perform well, but have larger  
96 confidence intervals around the estimates. However, the overall conclusions of previously published empirical  
97 data sets are not dependent on the method used.

## 98 Methods

### 99 Data Simulation

100 In order to investigate the performance of commonly used methods, we simulate body size observations  
101 from a bounded power law distribution using the inverse method, as described in Edwards et al. (2017)(2017).  
102 Let  $M$  be a random variable of body sizes described by the probability density function:

$$f(M) = CM^\lambda, M_{min} \leq M \leq M_{max}$$

103  $M_{min} = 0.0026$  and  $M_{max} = 1.2 * 10^3$ . These values are based on empirical body sizes of stream benthic  
104 communities reported in Pomeranz et al. (2020)(Pomeranz et al. 2022). Our results are not dependent on  
105 the range of body sizes, and we show the results of other body size ranges in the supplemental information.

106 For the main analysis, presented here, we sampled  $n = 1000$  body sizes from distributions described by  
107 five different  $\lambda$ 's: (-1.5, -1.75, -2.00, -2.25, -2.5), such that our model for  $\lambda$  is:

$$\lambda = Norm(\alpha - 0.5X, \sigma)$$

108 Each value of  $\lambda$  was assumed to come from a community across a hypothetical environmental gradient  
109  $X$ . Values of  $X$  for each community are uniformly distributed from -1 to 1. We repeated the data simulation  
110 1000 times (reps). The main results presented here were not dependent on the range of  $x$ -values or the  
111 number of sites (supplemental information).

### 112 Sample size, $n$

113 The number of observations in our simulations may bias the results. Therefore, we repeated the simulations  
114 described above, but varied the sample size  $n$ . We tested values of  $n = 200, 500, 1000, 5000, 10000$ .

### 115 “No” relationship

116 It is possible that our sampling or simulation framework could bias the results obtained. In order to  
117 assess whether or not our simulations are robust to this, we performed independent random samples from  
118 a bounded power law for five communities across a hypothetical gradient as described above, but we set  
119  $\lambda = -2$  for all five communities. Hence, we are simulating a scenario where the size spectra relationship  
120 is invariant to our hypothetical environmental gradient, and we should expect the estimated relationship  
121 coefficient  $\beta_1 = 0$ .

### 122 Finer resolution of $\lambda$ values

123 For the main analysis, we were interested in determining whether or not the methods were able to  
124 detect relatively coarse changes to size spectra parameters, specifically, a change of 0.5 units across each  
125 value of  $X$ , or an absolute change in  $\lambda$  values of 1.0 units. However, many seminal works on variation  
126 in size spectra relationships in freshwater are on a much finer scale (on the order of 0.1 to 0.25 units).  
127 Therefore, we performed simulations as described above, but varied  $\lambda$  from -1.9 to -2.1 across the hypothetical  
128 environmental gradient. This equates to a know relationship coefficient of  $\beta_1 = 0.1$  across the hypothetical  
129 gradient, and an absolute change of 0.2 units.

130 **Estimation of Size spectra parameter  $\lambda$**

131 For each sample of  $M$  and each of the 1000 replicates, we estimated the exponent  $\lambda$  using MLE methods  
132 modified from the the `sizeSpectra` package (Edwards et al. 2020). In addition, we use two common binning  
133 methods to estimate the OLS slope parameter in log-log space. For the first binning method, we created 6  
134 equal logarithmic bins covering the range of body sizes. The count in each bin was normalized by dividing  
135 by the bin width. This method has been used by (refs here, I think a lot of Woodward's group uses this...)  
136 Throughout the manuscript, the normalized equal logarithmic binning method will be referred to as ELBn.  
137 The second method was similar to ELBn, but bins of  $\text{Log}_2$  widths are used. The count in each bin is  
138 normalized in the same way. This method has been used by (Pomeranz et al. 2019, McGarvey et al. 2019)  
139 and is referred to as the Normalized Abundance Spectrum (NAS). Previous work has investigated the biases  
140 in  $\lambda$  estimates when using different methods (Edwards et al. (2017)(2017), Edwards et al. 2020(2020), White  
141 et al. 2008(2007)). The focus of the present work is to investigate biases when estimating the relationship or  
142 relative change in parameters across an environmental gradient. Hence, we do not discuss the  $\lambda$  estimates  
143 in the main text, but the results are presented in the Supplmental Information.

144 **Estimation of relationship across gradient,  $\beta_1$**

145 For each simulation replicate, we estimated how the size spectra parameters varied across the hypothetical  
146 gradient. Simple OLS regression were conducted in the form  $\lambda_{\text{estimate}} \beta_0 + \beta_1 * X$ . The distribution of the  
147 relationship coefficient,  $\beta_1$ , were plotted compared to the known relationship.

148 **Empirical Data**

149 We re-analyze two data sets of benthic macroinvertebrate communities from stream habitats across two  
150 different gradients. In the first, quantitative macroinvertebrate samples were collected from streams across  
151 an acid mine drainage (AMD) stress gradient. Details of the sample collection and processing can be found  
152 in (Pomeranz et al. 2019). Briefly, all individuals from each sample were identified to the lowest practical  
153 taxonomic unit and body lengths were measured using image processing software from photos taken with a  
154 camera mounted to a dissecting microscope. Body mass was estimated using taxon-specific published length  
155 weight regressions.

156 The second data set was from the wadeable stream sites of the National Ecological Observatory Network  
157 (National Ecological Observatory Network (NEON) 2022). NEON stream sites are located across a wide  
158 temperature gradient in the United States, from Puerto Rico to Alaska. Quantitative macroinvertebrate  
159 samples were collected using the most appropriate method based on the local habitat. All individuals were  
160 identified and had their body lengths measured, and body mass was estimated using published length weight  
161 regressions. This data has been analyzed previously using size spectra methods as described in (Pomeran-  
162 anz et al. 2022). Detailed methods of the sampling collection and data processing methods can be found  
163 on the NEON website ([<https://www.neonscience.org/data-collection/macroinvertebrates>], or macroinverte-  
164 brate DPI pubs). Coefficient estimates ( $\pm 1$  standard deviation) are compared across methods. This allows  
165 us to determine whether or not the main results would differ depending on the method used.

166 **Results**

167 **Relationship across the hypothetical environmental gradient**

168 Alternate plot version

169 All methods performed reasonably well in detecting the known relationship across the hypothetical en-  
170 vironmental gradient (mean  $\beta_1$  estimates ranged from -0.419 to -0.498). However, the SD for the NAS and  
171 ELBn methods were 2-3x larger than that of the MLE method (Figure xx). This result was consistent across  
172 simulation parameters, including sample size, number of sites, and  $\lambda$  values.

## PLB

$m_{\text{range}} = (0.0026, 1200)$

$b = (-2.5, -1.5)$

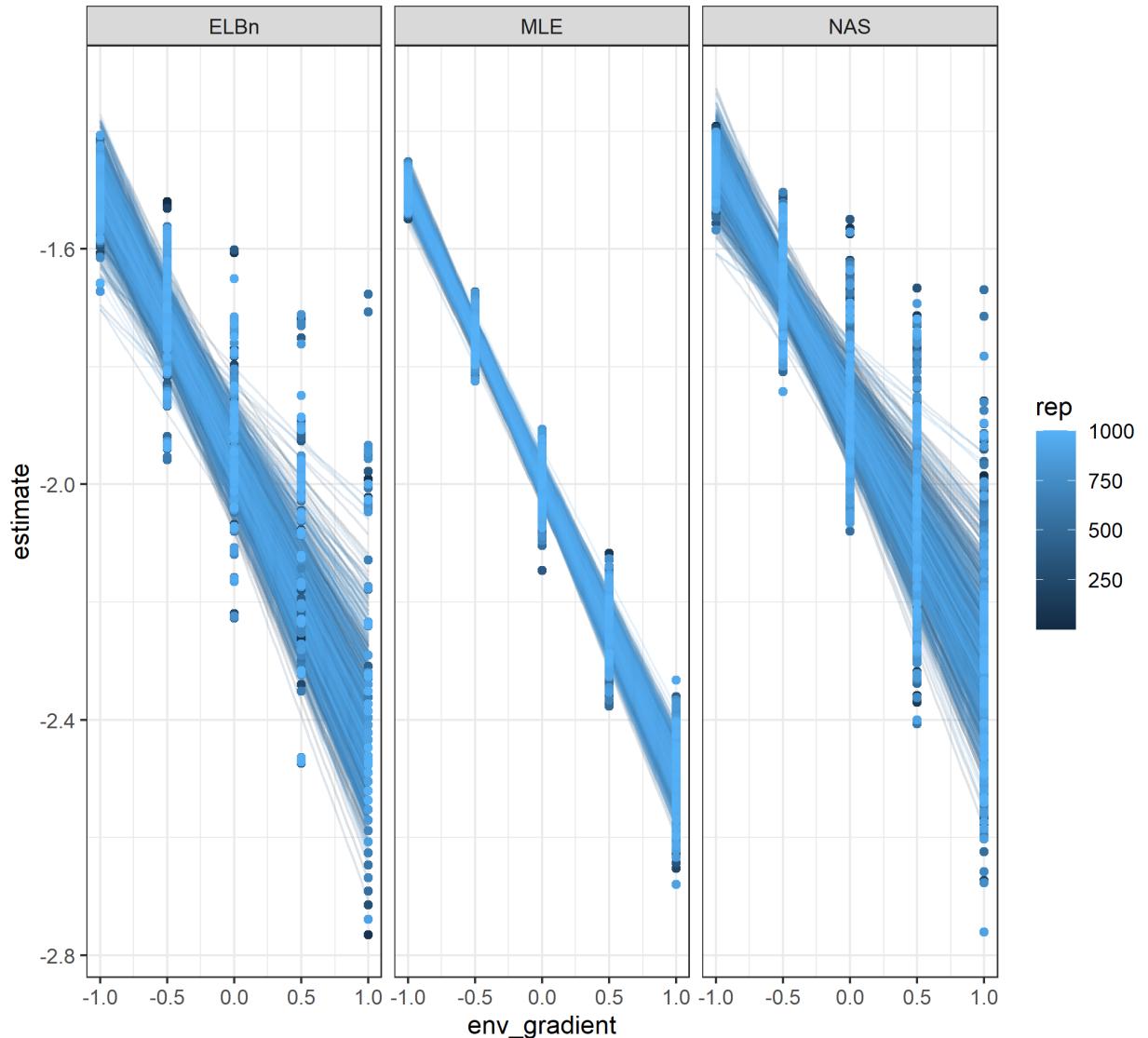


Figure 1: Relationship estimates across the hypothetical gradient for each replicate. Each panel is a different method for estimating the size spectra parameter.

## PLB

m\_range =(0.0026,1200)  
 $b = (-2.5, -1.5)$

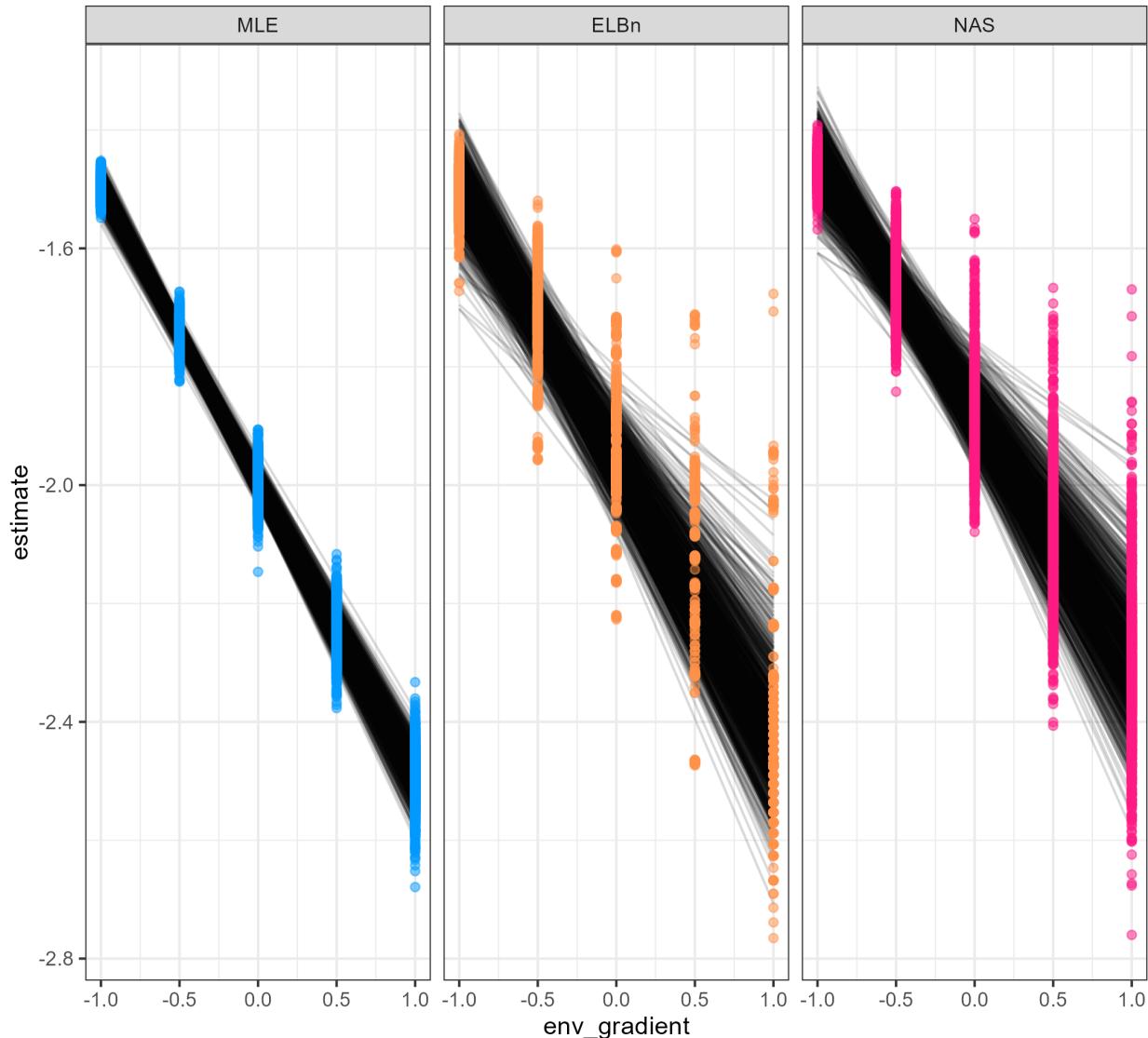


Figure 2: alternate plot version, thoughts on which you prefer??

```
173 ## Model p50 p25 p975  
174 ## 1 MLE -0.49 -0.50 -0.48  
175 ## 2 ELBn -0.49 -0.52 -0.29  
176 ## 3 NAS -0.44 -0.46 -0.30
```

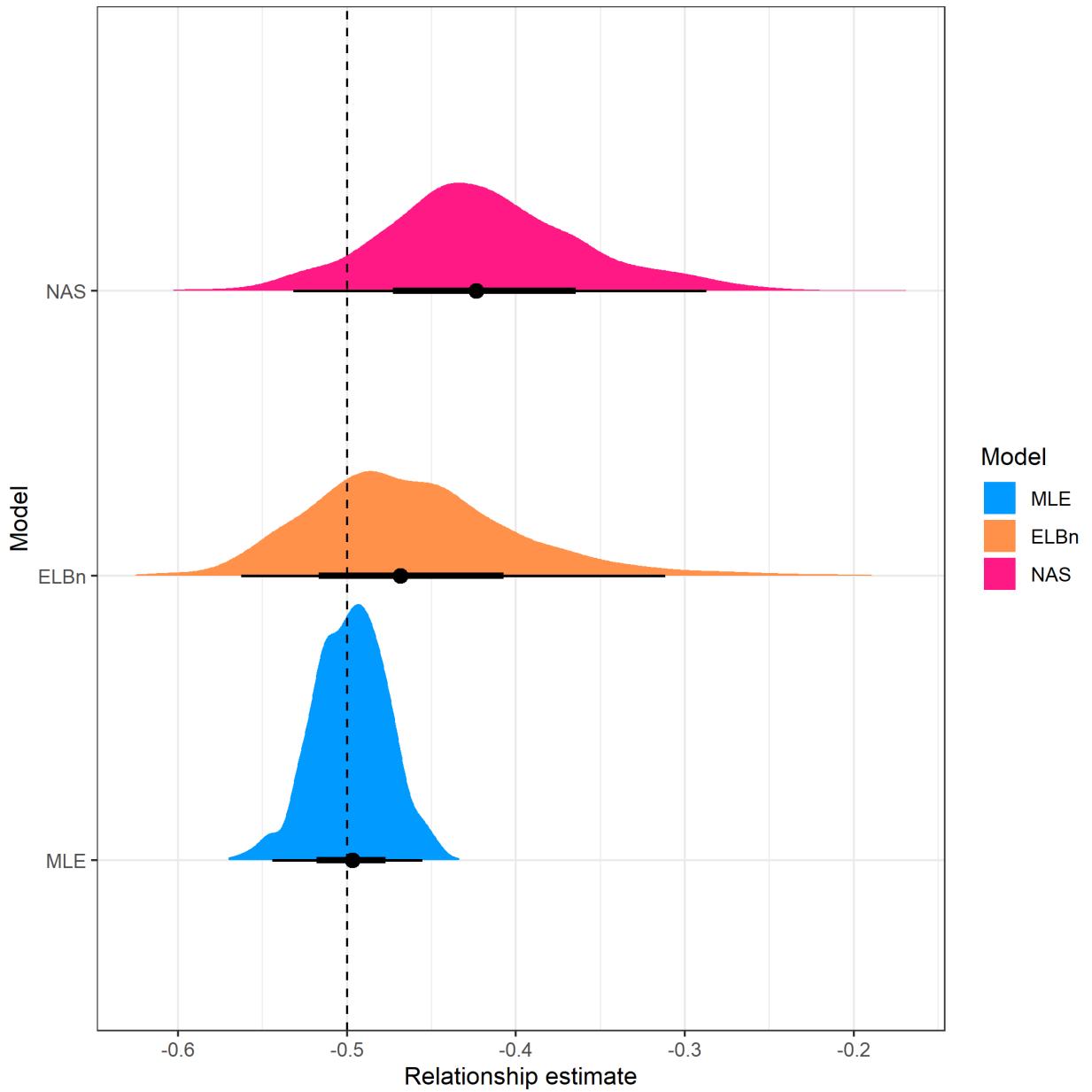


Figure 3: Halfeye plots for the estimated relationship coefficient estimates across 1000 replicates and three different methods. The point below each distribution plot is the median estimate, the thick and thin bars represent the 66th and 95th percentile, respectively. Vertical line is the known relationship. The median estimate for the MLE method was closest to the known relationship, and had the narrowest distribution across replicates.

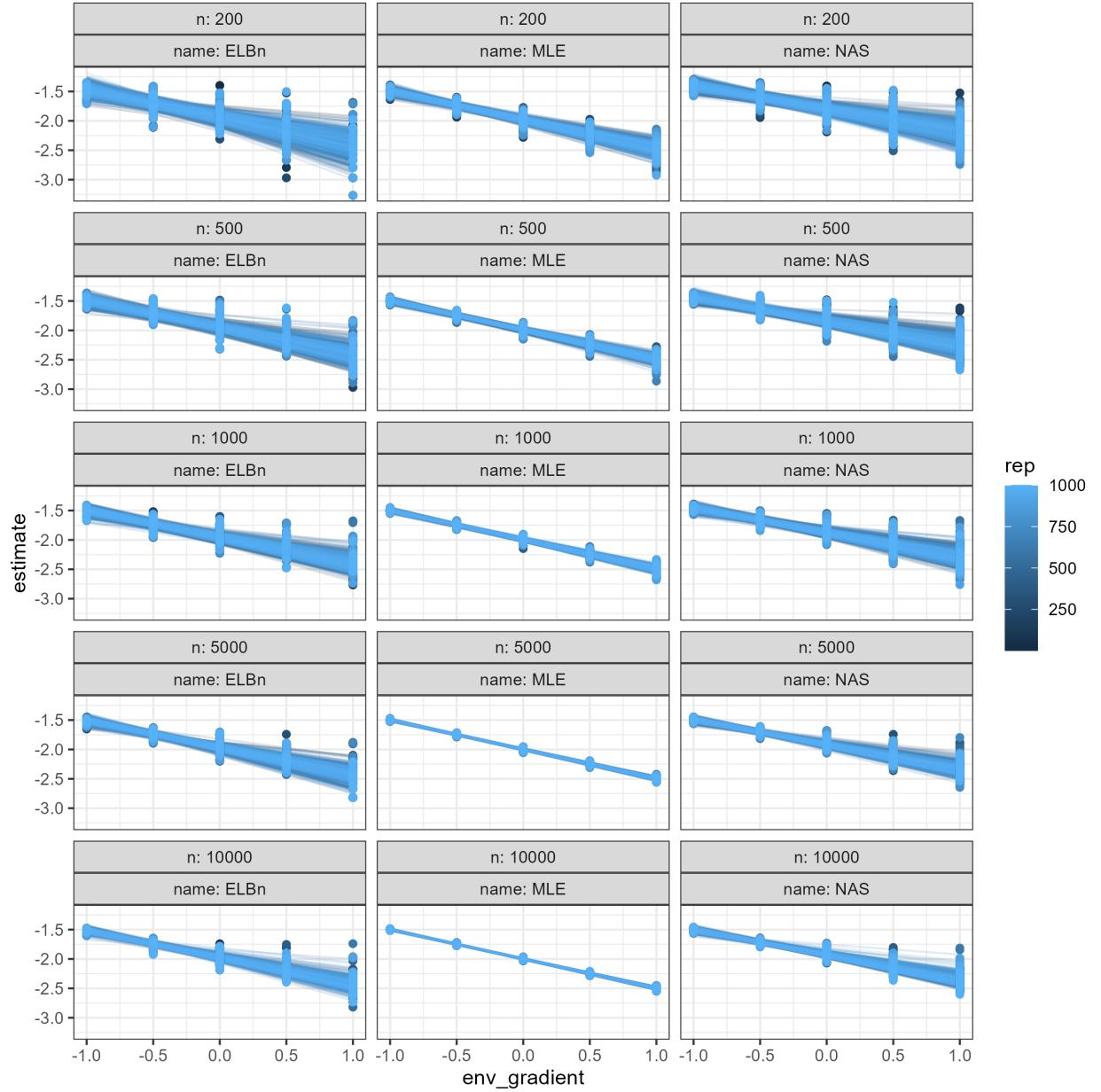


Figure 4: Individual regression estimates across the hypothetical gradient based on sample size (rows) and methodology used (columns).

177 **No relationship**

178 All of the methods performed similarly when there was no relationship across the hypothetical gradient.  
179 The mean  $\beta_1$  estimates for all three methods ranged from -0.0005 to 0.0002. However, the distribution  
180 of coefficient estimates was much wider for the two binning methods; the 95th percentile of the MLE  
181 distributions is narrower than the 66th percentiles for the NAS and ELBn methods.

```
182 ## Model p50 p25 p975
183 ## 1 MLE 0 -0.01 0.04
184 ## 2 NAS 0 -0.04 0.12
185 ## 3 ELBn 0 -0.05 0.13
```

186 **Small variation in lambda**

187 The performance of all methods to detect small variation in the  $\lambda$  parameter (between -1.9 to -2.1) across  
188 the environmental gradient was also investigated. The median  $\beta_1$  coefficient estimates for the MLE method  
189 was -0.1, which exactly matched the value for the hypothetical relationship, while the other two methods  
190 both had median estimates of 0.9. The upper 97.5 percentile of the distribution was below 0 for the MLE  
191 method, meaning that most of the estimates had the correct sign. In contrast, the upper 97.5 percentile for  
192 both the NAS and ELBn methods were greater than 0 at 0.027 and 0.052, respectively. (Figure 4).

```
193 ## Model p50 p25 p975
194 ## 1 MLE -0.10 -0.11 -0.06
195 ## 2 NAS -0.09 -0.13 0.03
196 ## 3 ELBn -0.09 -0.13 0.05
```

197 **Empirical data**

198 For both empirical data sets, the direction and magnitude of change (i.e.  $\beta_1$  coefficients) are generally in  
199 agreement. Size spectra parameters consistently increase (become flatter) in the AMD data (Fig. XX A).  
200 Likewise, the size spectra parameters consistently increase (become steeper) with increasing temperature  
201 across the NEON sites (Fig XX D)

202 **NOTE** I need to re-plot these in ggplot and stitch together. Currently stitching pre-made png files,  
203 I think that's why it looks so funky.

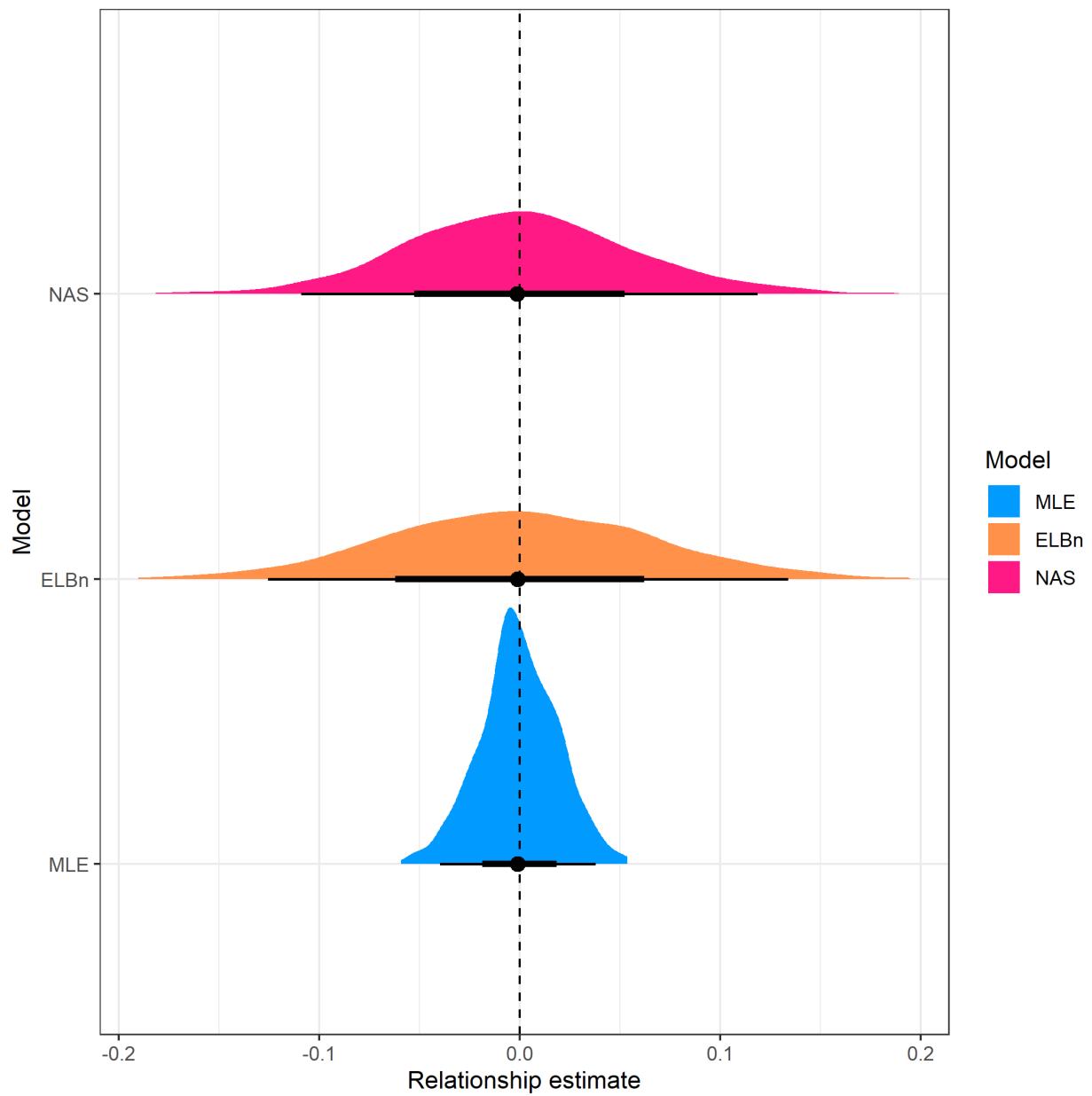


Figure 5: Individual regression estimates when no relationship exists across the hypothetical environmental gradient. The methods are separated by rows. The median for all methods is approximately zero, but the distribution for the MLE method is narrower than the two binning methods.

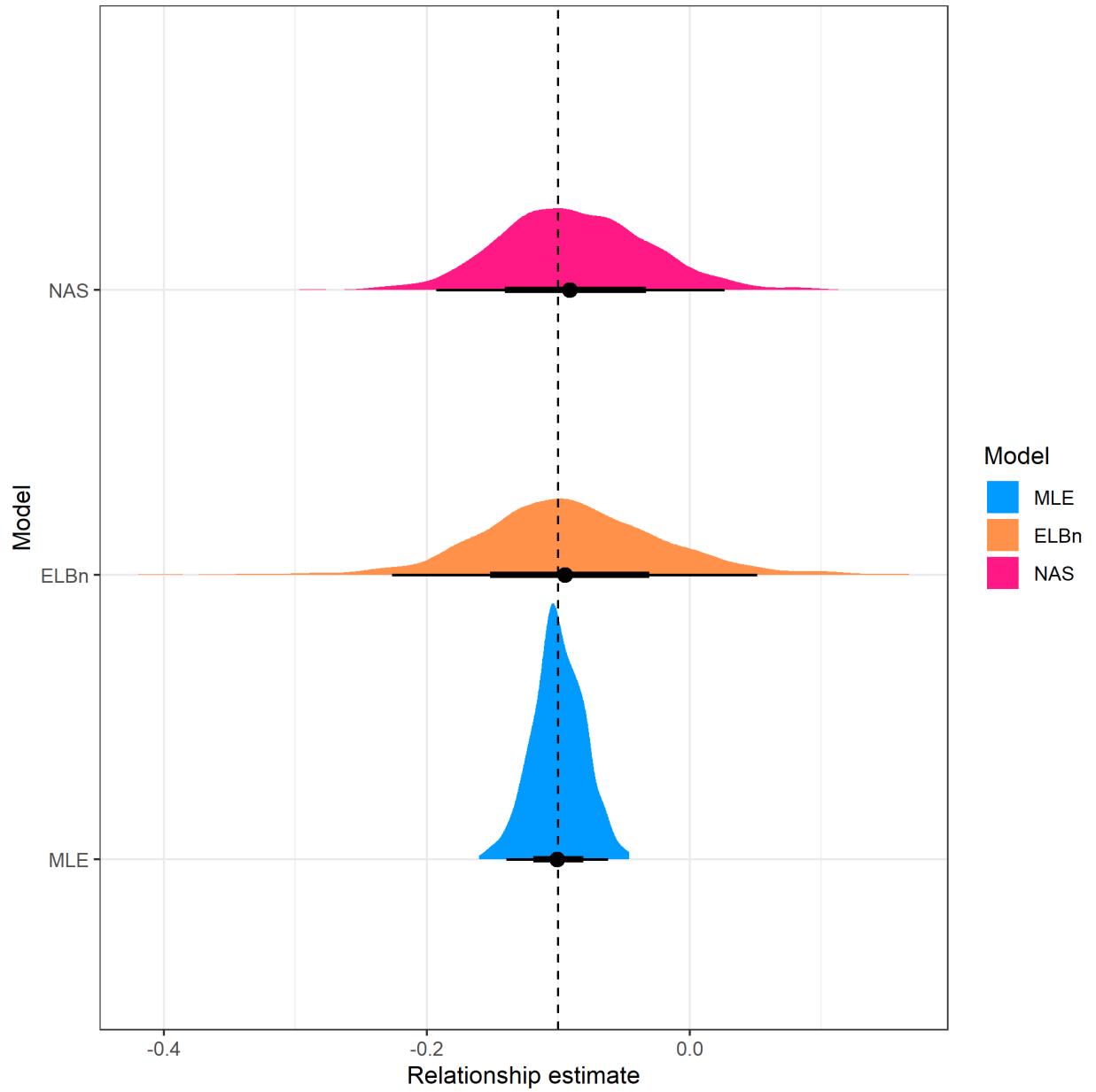


Figure 6: Distribution of the estimated relationship for each method. Here,  $\lambda$  varied from -1.9 to -2.1, with a known relationship of -0.1 across the gradient indicated by the dashed vertical line. The distribution of relationship estimates was approximately centered at the known relationship for all methods, with a narrower distribution observed for the MLE estimates.

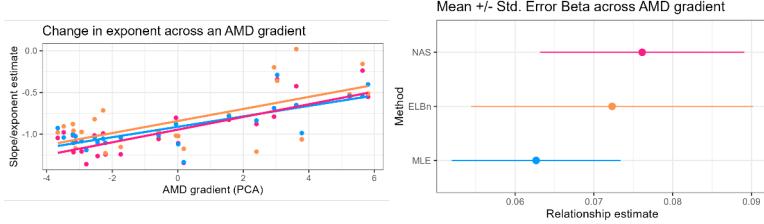
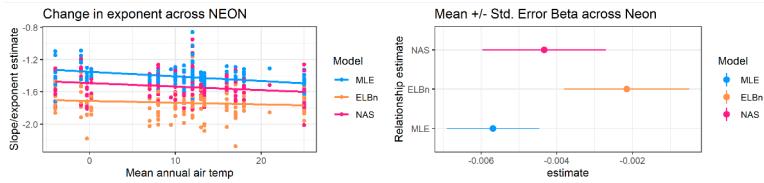
**A****B****C****D**

Figure 7: Figure of empirical data estimates. All of the methods estimate the same sign of the relationship, but the estimates from the binning methods are generally greater than the MLE estimates.

204 The  $\beta_1$  coefficient estimates for the AMD data are all positive and of relatively similar magnitudes across  
205 methods. The MLE estimate has the smallest magnitude and SD ( $0.063 \pm 0.011$  SD), The ELBn method is  
206 slightly larger and has a wider SD : ( $0.072 \pm 0.018$ ). Finally, the NAS method has the largest estimate and  
207 an intermediate standard deviation ( $0.076 \pm 0.013$ ). The range of the gradient in the AMD data is relatively  
208 small (9.5), meaning that the absolute change in the size spectra parameter ( $\lambda$ ) across the AMD gradient  
209 are similar regardless of method used (range of absolute: 0.59 to 0.72).

210 Similar to the AMD results, the  $\beta_1$  coefficient estimates for the NEON data are all negative and have  
211 similar magnitudes across methods. The MLE estimate is once again the smallest ( $-0.006 \pm 0.001$  SD).  
212 The NAS and ELBn methods are larger at  $-0.004 \pm 0.002$ , and  $-0.002 \pm 0.002$ , respectively. Although the  
213 estimates are similar and have a small magnitude, the range of the gradient in this data is larger (~30°C),  
214 meaning that the absolute change in  $\lambda$  estimates ranges from 0.06 to 0.17, depending on the method used.

## 215 Simulating shallower lambdas

216 Estimates of the relationship between lambda and a hypothetical gradient varied depending on the method  
217 used. However, interpretations of empirical data were broadly consistent across methodologies. Upon closer  
218 inspection, the values of size spectra parameters in the empirical data were considerably shallower than -2.  
219 Given that we found the performance of all methods increased with shallower  $\lambda$ 's we wanted to investigate  
220 how the methods performed with simulated values closer to the empirical estimates of parameters describing  
221 size spectra relationships. Therefore, we repeated the simulation process as in the main analysis, but used  
222  $\lambda$  values ranging from -1.1 to -1.5, with a known relationship of  $\beta_1 = -0.2$

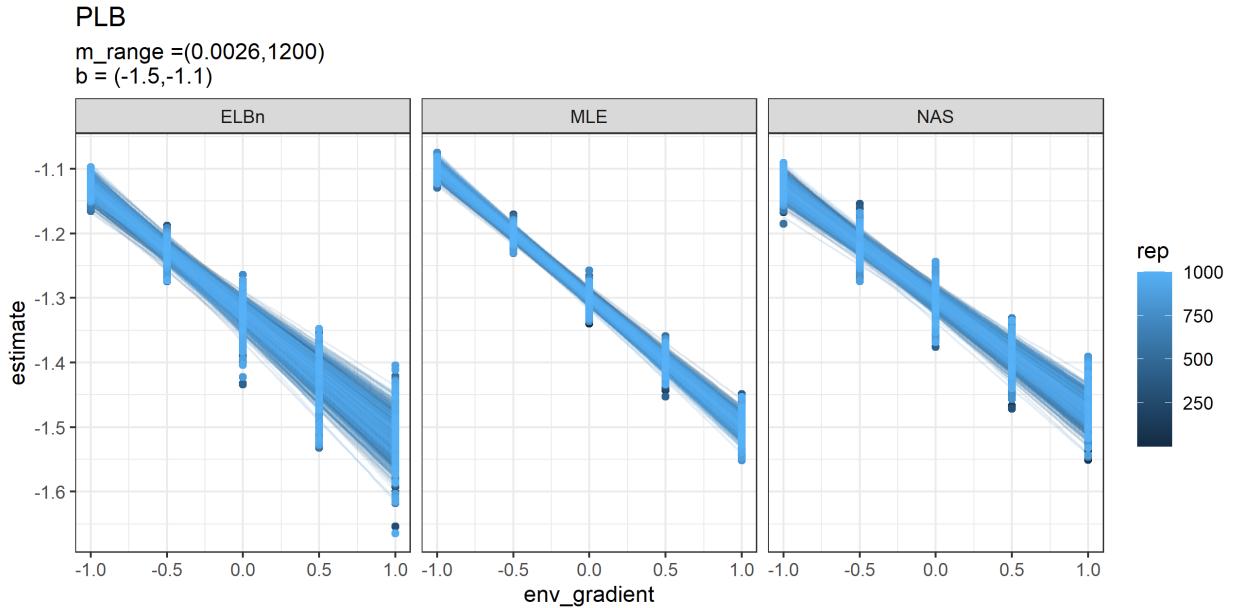


Figure 8: Relationship estimates across the hypothetical gradient for each replicate. Each panel is a different method for estimating the size spectra parameter.

223 We found generally less variability in the relationship estimates ( $\beta_1$ ) across a gradient of distributions  
 224 described with shallower  $\lambda$  parameters. However, the median MLE estimate was once again closest to the  
 225 hypothetical relationship and had the narrowest distribution. The median ELBn estimate was close to the  
 226 hypothetical relationship but had a slightly wider distribution. Finally, 95% of the NAS estimates were  
 227 greater than the hypothetical relationship. In general, the distributions from the different methodologies  
 228 followed the same overall pattern, with the MLE estimates being the smallest, and the binning methods  
 229 being larger (i.e., more positive).

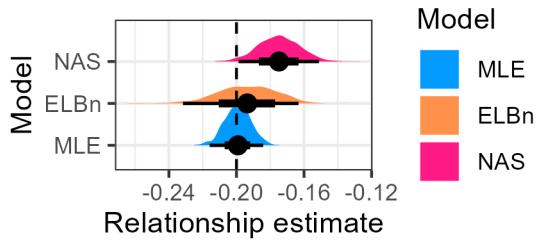


Figure 9: Distribution of relationship coefficient estimates. Vertical line is the known relationship. All methods under estimate the value, but the mean magnitude and distribution of values is greater for the ELBn and NAS methods.

```
230 ##   Model   p50   p25   p975
231 ## 1   MLE -0.20 -0.21 -0.18
232 ## 2   ELBn -0.19 -0.21 -0.16
233 ## 3   NAS -0.17 -0.18 -0.15
```

234 **Discussion**

235 Measuring parameters describing the decline in abundance with increasing body size in communities  
236 is being done with increasing frequency across ecology. Previous work has investigated the accuracy and  
237 inherent biases associated with different estimation methods. However, how these inaccuracies and biases  
238 compound across environmental gradients remains uncertain, making it difficult to detect variation in size  
239 spectra parameters across environmental gradients with confidence. Here, we sampled body sizes from known  
240 distributions with varying parameters ( $\lambda$ ) and estimated the coefficient of the relationship ( $\beta_1$ ) across a hy-  
241 pothetical gradient in order to assess how the results would vary depndng on the method used. Likewise, we  
242 compared how the interpretation of previously published results could change depending on the methodology  
243 used to estimate the relationship.

244 Generally, the estimate from the MLE method was always closer to the known relationship in simulated  
245 data, and always had the smallest variation around the estimate.

246 Binning methods are easy to use and interpret, which most likely accounts for their wide use in ecological  
247 studies. However, aggregating individuals into logarithmic bins removes a large source of the variation within  
248 the data by collapsing body size variation into a single values within bins. For example, all individuals placed  
249 into a bin that ranges from 2-4 grams of mass are all treated as having a mass of 3 grams, the midpoint  
250 of that bin. Likewise, a single abundance value is taken for this bin, despite that fact that there is almost  
251 certainly variation in the abundance of individuals that weigh ~2 grams or ~4 grams. By homogenizing  
252 the data in this way, it seems that binning methods will always produce noisier results than MLE from the  
253 binning process alone, since binning is akin to deleting information that the model could otherwise use.

254 One of the benefits of using MLE is that all data points are retained for use within the model.

255 The performance of all methods improved as sample size increases, and as the exponent of the size  
256 spectra relationship gets larger (more shallow). As either or both of these variables change, the difference  
257 in estimated relationship coefficients declines across the methods. This is particularly interesting given the  
258 fact that both empirical data sets of stream communities examined here are shallower than the expected  
259 relationship of  $\lambda = -2$ , and the conclusions of the change in size spectra relationships across the gradients  
260 are not dependent on method used. If communities under study are in fact described by shallower exponents,  
261 the method used may not be critical in the conclusions reached. However, the MLE method performs as  
262 well or better than the two binning methods examined here under all contexts and should be the preferred  
263 method in future studies. At a minimum, future studies should report MLE estimates to ensure that the  
264 results are not dependent on the methodology used.

265 Despite the drop in performance with reduced variation in  $\lambda$  values, when a significant relationship was  
266 observed, it was generally in the correct direction and of a similar magnitude. This suggests that previously  
267 reported significant changes in size spectra parameters across environmental gradients and in experimental  
268 manipulations are plausible, and the magnitude of the relative change is a reasonable estimate. Given that  
269 all of the data within a study is treated identically, the the over all change in size spectra parameters is  
270 likely reasonable. However, the biases and inconsistencies in relationship estimates presented here suggest  
271 that it would be difficult if not impossible to compare the relative changes across different published studies  
272 which use different methods. Publication of individual body size data with future studies of size spectra  
273 relationships would greatly aid in our ability to generalize changes to this fundamental aspect of community  
274 organization across spatiotemporal scales and in response to environmental stressors and perturbations.

275 **Concluding Remarks**

276 Both of the binning methods can perform as well as MLE under certain circumstances (large sample size,  
277 shallow  $\lambda$ ). However, it would be extremely difficult to know if this was the case beforehand. Furthermore,  
278 in order to confirm this result, MLE methods would likely need to be employed. With the publication of the  
279 sizespectra package (2017), producing MLE estimates of size spectra parameters is a relatively easy task.  
280 Therefore, we recommend using this in all future studies of size spectra relationships.

281 We reiterate the recommendations of White et al. (2007)(2007), Sprules and Barth (2016)(2016) and  
282 Edwards et al. (2017)(2017) to estimate size spectra relationships using MLE methods due to their superior  
283 performance in nearly every context. Furthermore, we strongly encourage authors to publish individual size  
284 data whenever possible. This will allow for the consistent re-analysis of existing data sets as methodologies

<sup>285</sup> develop and improve. This will aid in the ability for size spectra work to be synthesized between research  
<sup>286</sup> groups and across scales.

287 **References**

- 288 Edwards, A. M., J. P. W. Robinson, J. L. Blanchard, J. K. Baum, and M. J. Plank. 2020. Accounting for the  
289 bin structure of data removes bias when fitting size spectra. *Marine Ecology Progress Series* 636:19–33.
- 290 Edwards, A. M., J. P. W. Robinson, M. J. Plank, J. K. Baum, and J. L. Blanchard. 2017. Testing and  
291 recommending methods for fitting size spectra to data. *Methods in Ecology and Evolution* 8:57–67.
- 292 McGarvey, D. J., T. E. Woods, and A. J. Kirk. 2019. Modeling the Size Spectrum for Macroinvertebrates  
293 and Fishes in Stream Ecosystems. *Journal of Visualized Experiments: JoVE*.
- 294 National Ecological Observatory Network (NEON). 2022. Macroinvertebrate collection (DP1.20120.001).  
295 National Ecological Observatory Network (NEON).
- 296 Petchey, O. L., and A. Belgrano. 2010. Body-size distributions and size-spectra: Universal indicators of  
297 ecological status? *Biology Letters* 6:434–437.
- 298 Pomeranz, J. P. F., J. R. Junker, and J. S. Wesner. 2022. Individual size distributions across North American  
299 streams vary with local temperature. *Global Change Biology* 28:848–858.
- 300 Pomeranz, J. P. F., H. J. Warburton, and J. S. Harding. 2019. Anthropogenic mining alters macroinverte-  
301 brate size spectra in streams. *Freshwater Biology* 64:81–92.
- 302 Sprules, W. G., and L. E. Barth. 2016. Surfing the biomass size spectrum: Some remarks on history, theory,  
303 and application. *Canadian Journal of Fisheries and Aquatic Sciences* 73:477–495.
- 304 White, E. P., S. K. M. Ernest, A. J. Kerkhoff, and B. J. Enquist. 2007. Relationships between body size  
305 and abundance in ecology. *Trends in Ecology & Evolution* 22:323–330.

306 **Supplementary material**

307 **Range of body sizes,  $M$**

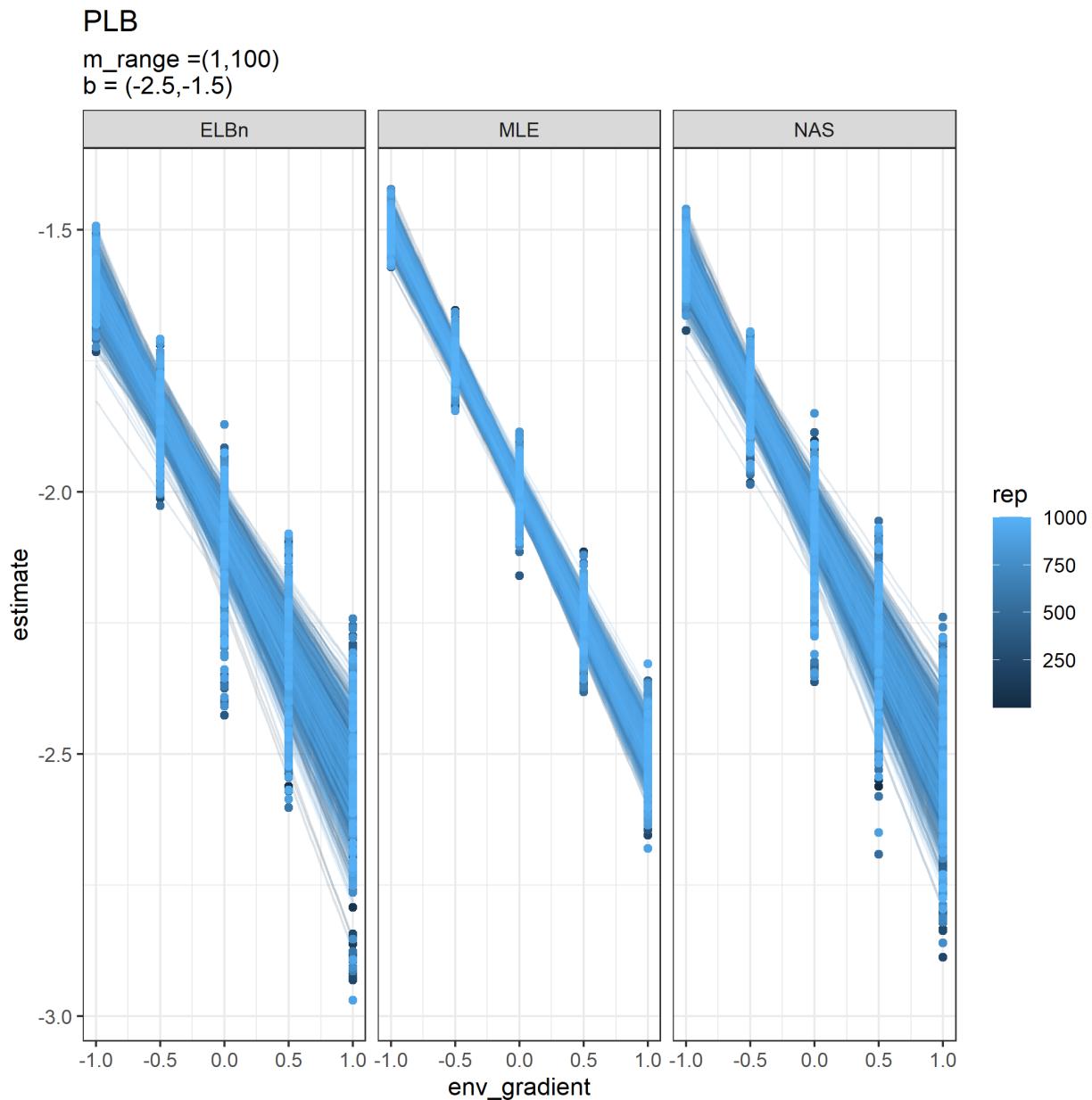


Figure S1: Individual regressions for five sites across a hypothetical gradient with a known relationship of 0.5. Range of body sizes is reduced and is from 1, to 100.

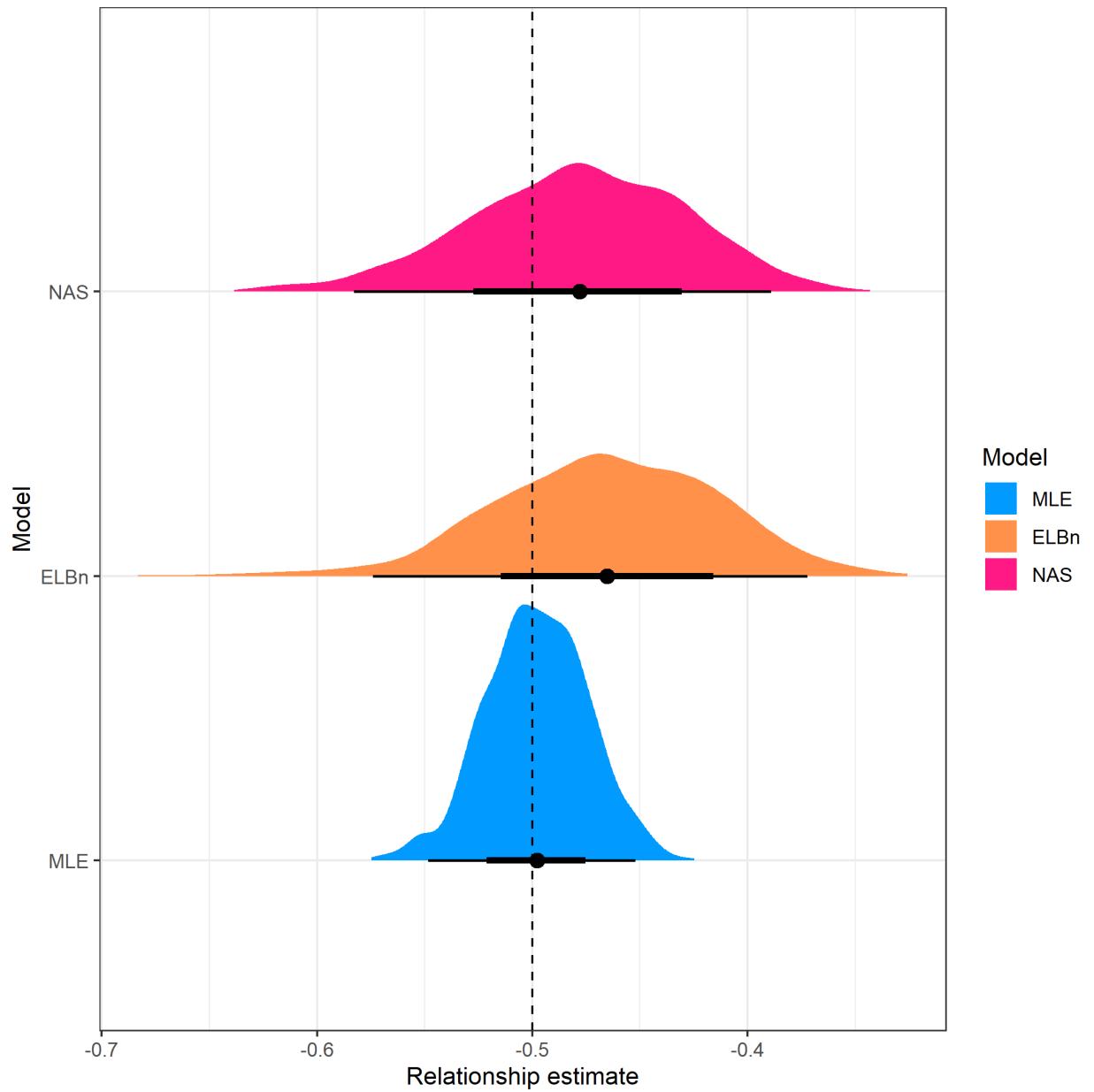


Figure S2: Distribution of estimated relationship ( $\beta_1$ ) coefficient's for five sites across a hypothetical gradient with known value of 0.5. Range of body sizes is reduced and is from 1, to 100.

308 Large environmental gradient

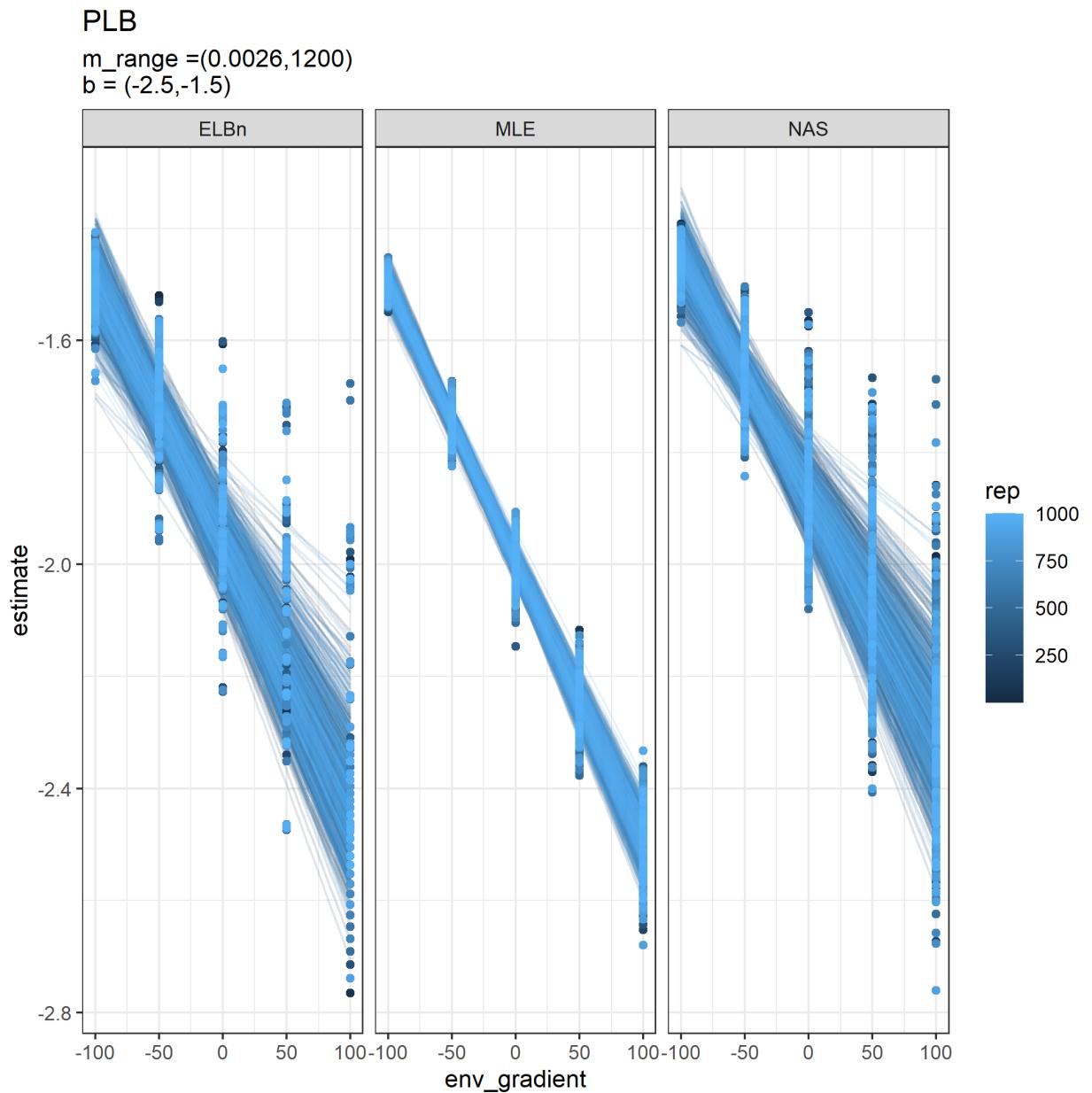


Figure S3: Individual regressions for five sites across a hypothetical gradient with a known relationship of 0.5. Range of environmental values ( $x$ -axis) increased to be -1000, to 1000.

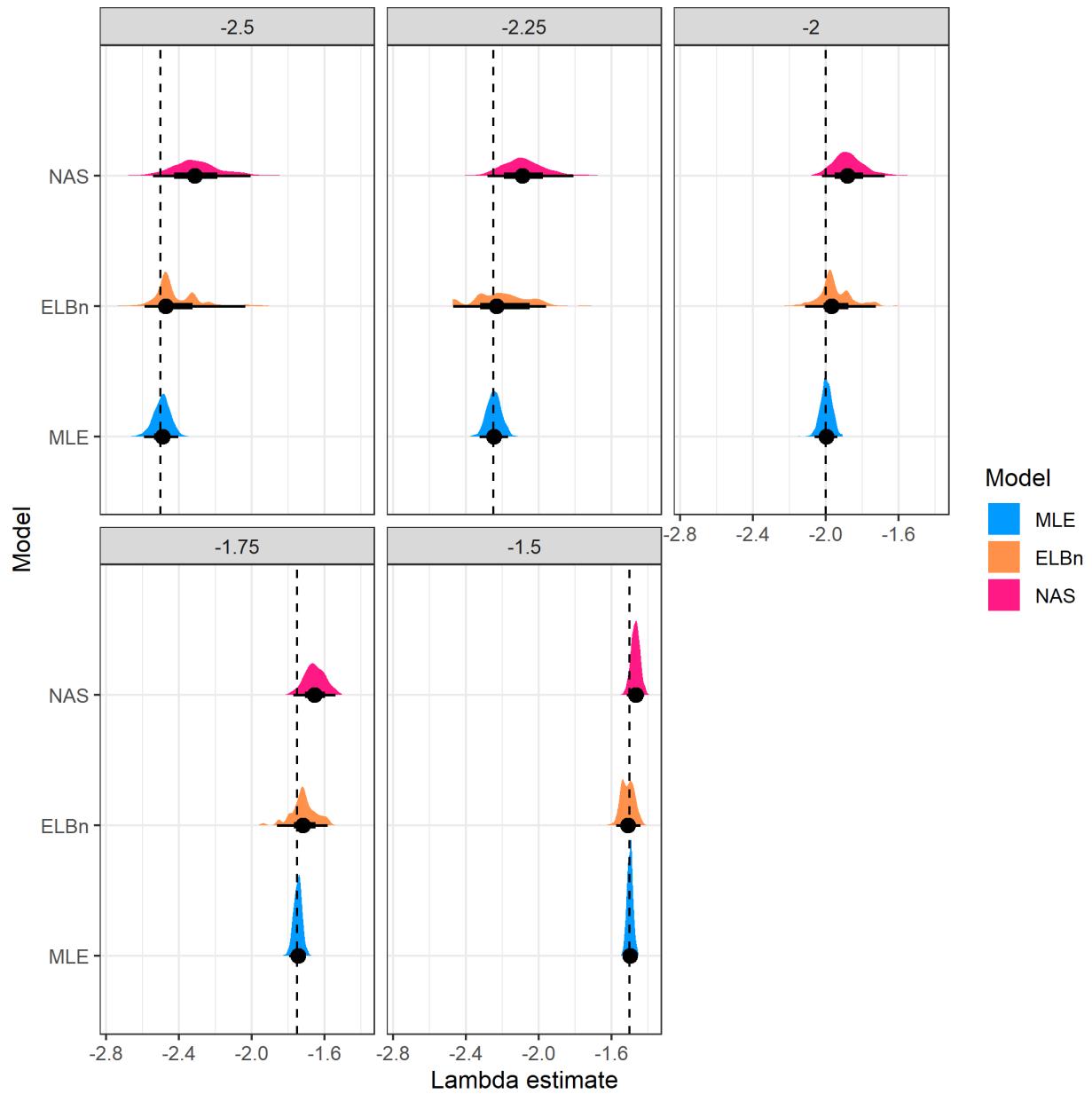


Figure S4: Distribution of estimated  $\lambda$  coefficient for five sites across a hypothetical gradient with known values. Range of environmental values ( $x$ -axis) increased to be -1000, to 1000.

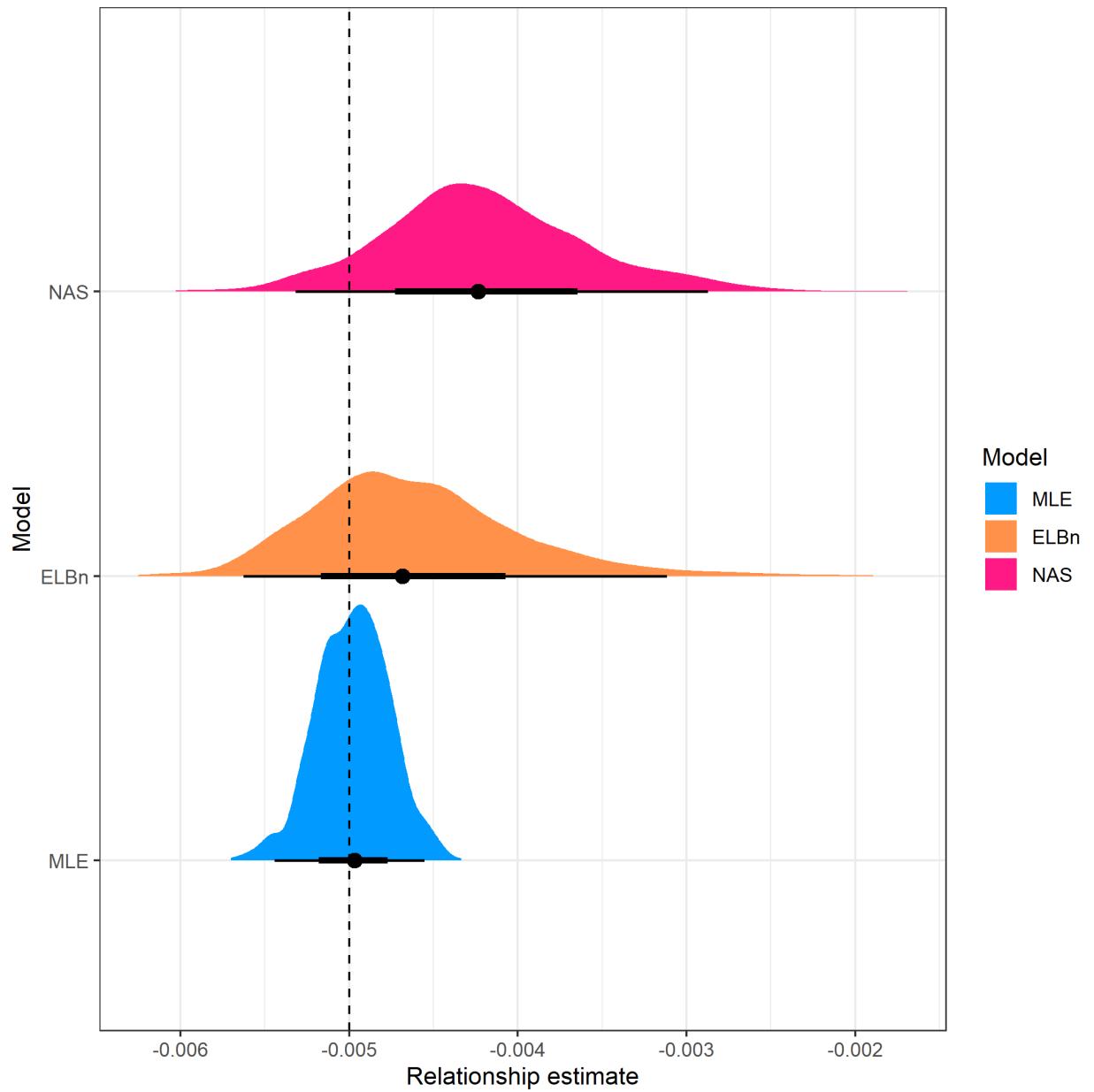


Figure S5: Distribution of estimated relationship ( $\beta_1$ ) coefficient's for five sites across a hypothetical gradient with known value of 0.5. Range of environmental values ( $x$ -axis) increased to be -1000, to 1000.

309 Varying number of sites

310 10 sites

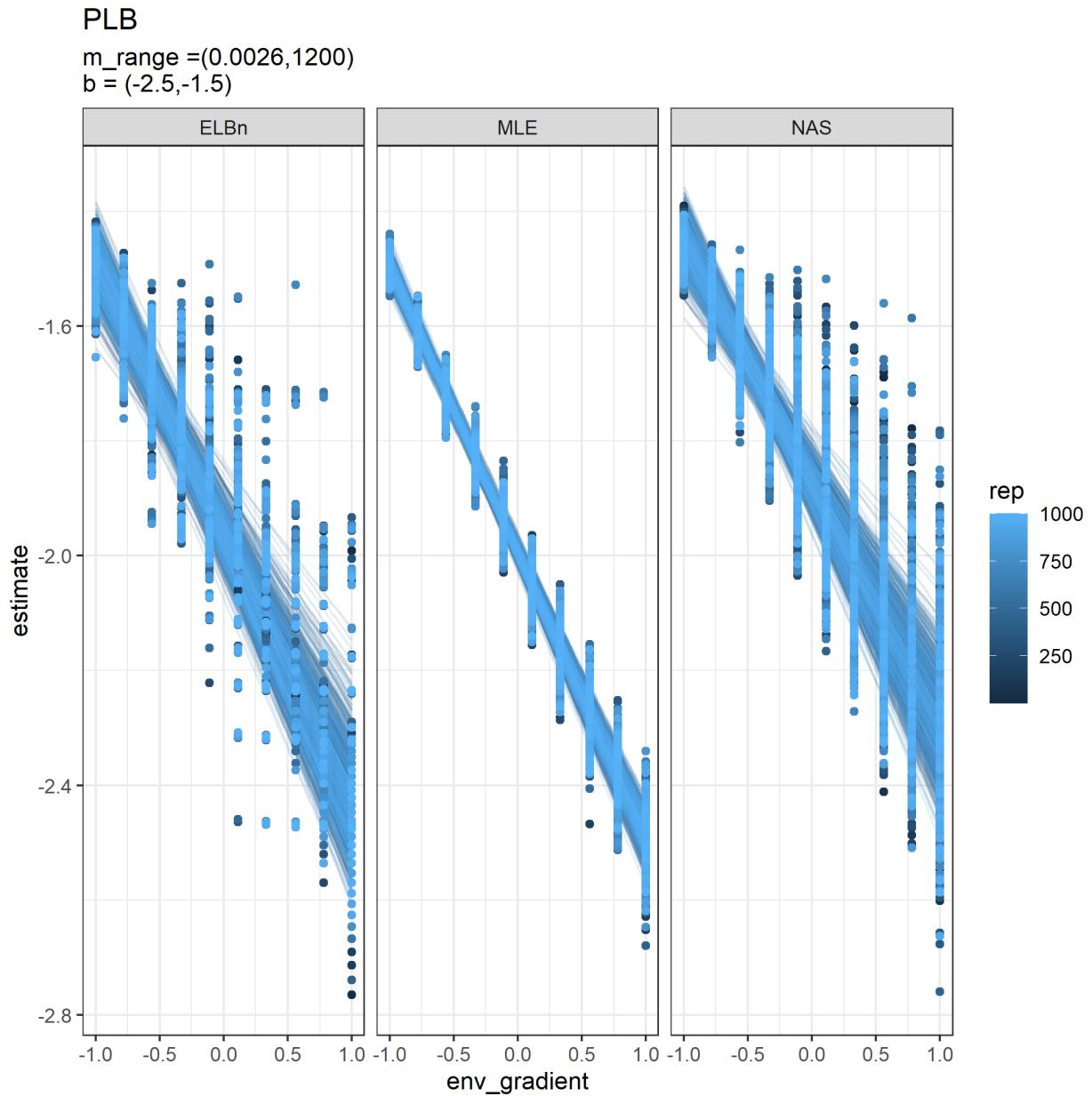


Figure S6: Individual regressions for ten sites across a hypothetical gradient with a known relationship of 0.5. All other parameters are the same as in the main analysis

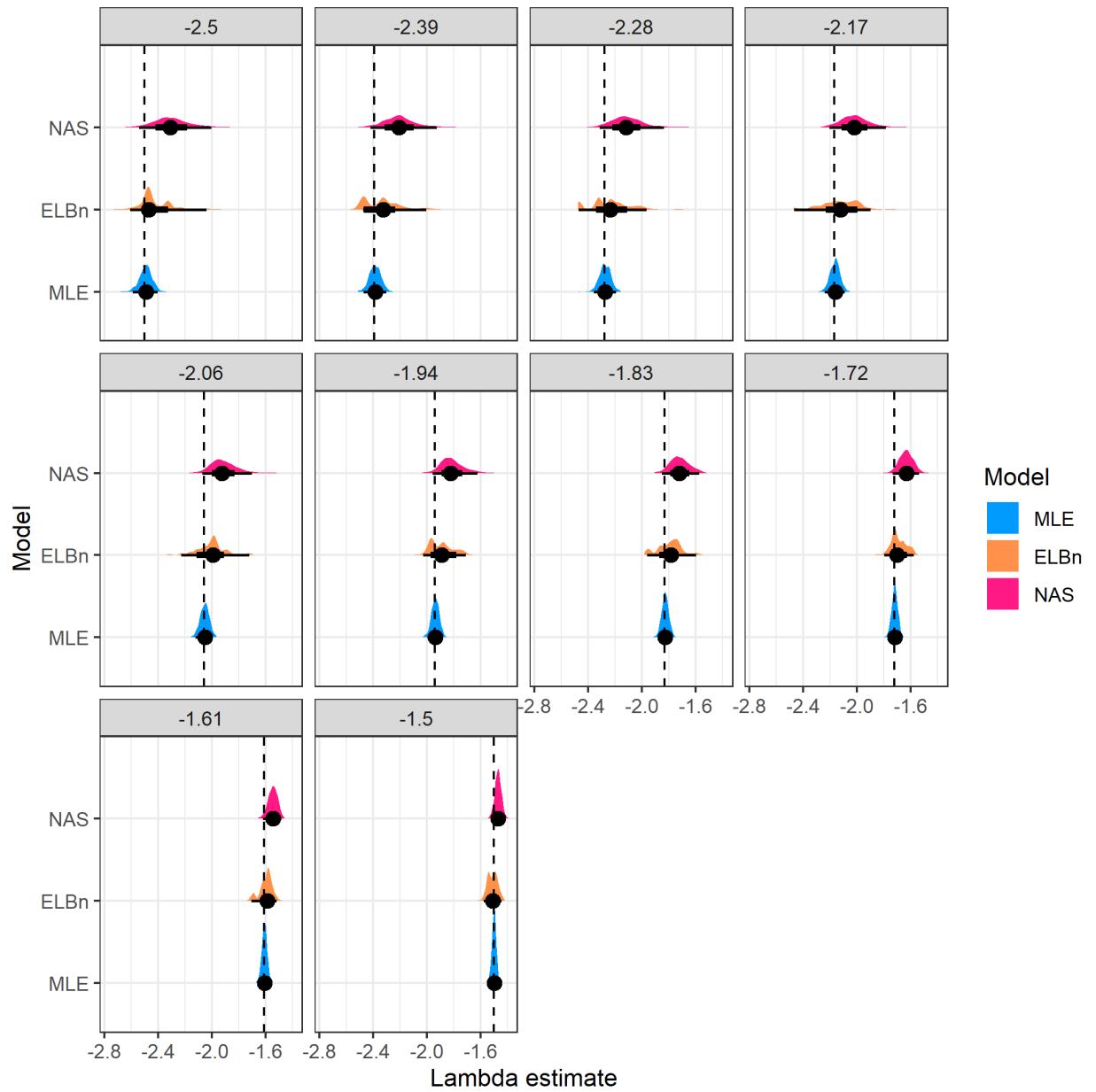


Figure S7: Distribution of estimated  $\lambda$  coefficient for ten sites across a hypothetical gradient with known values. All other parameters are the same as in the main analysis

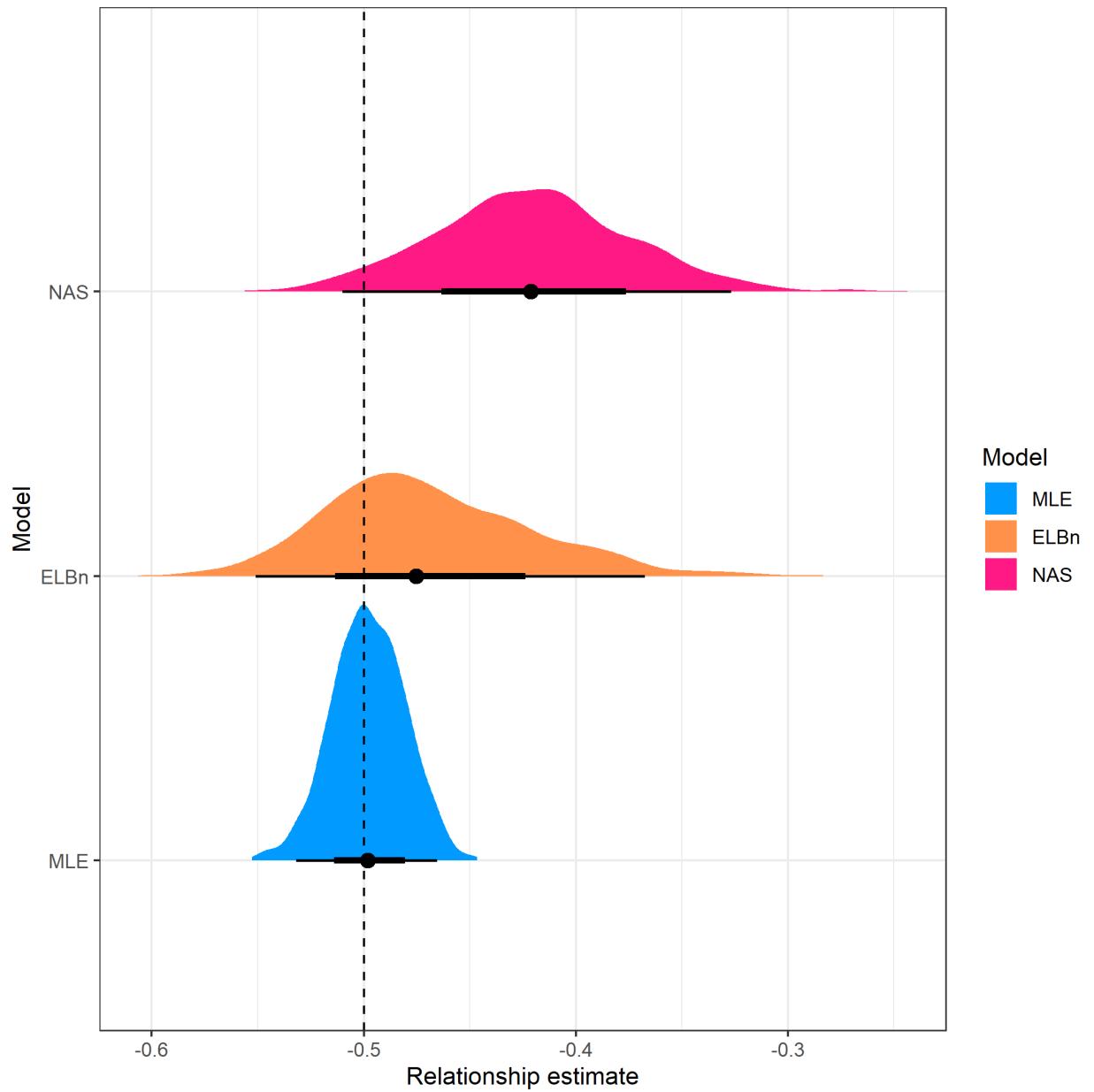


Figure S8: Distribution of estimated relationship ( $\beta_1$ ) coefficient's for ten sites across a hypothetical gradient with known value of 0.5. All other parameters are the same as in the main analysis

311 Three sites

PLB

m\_range =(0.0026,1200)

b = (-2.5,-1.5)

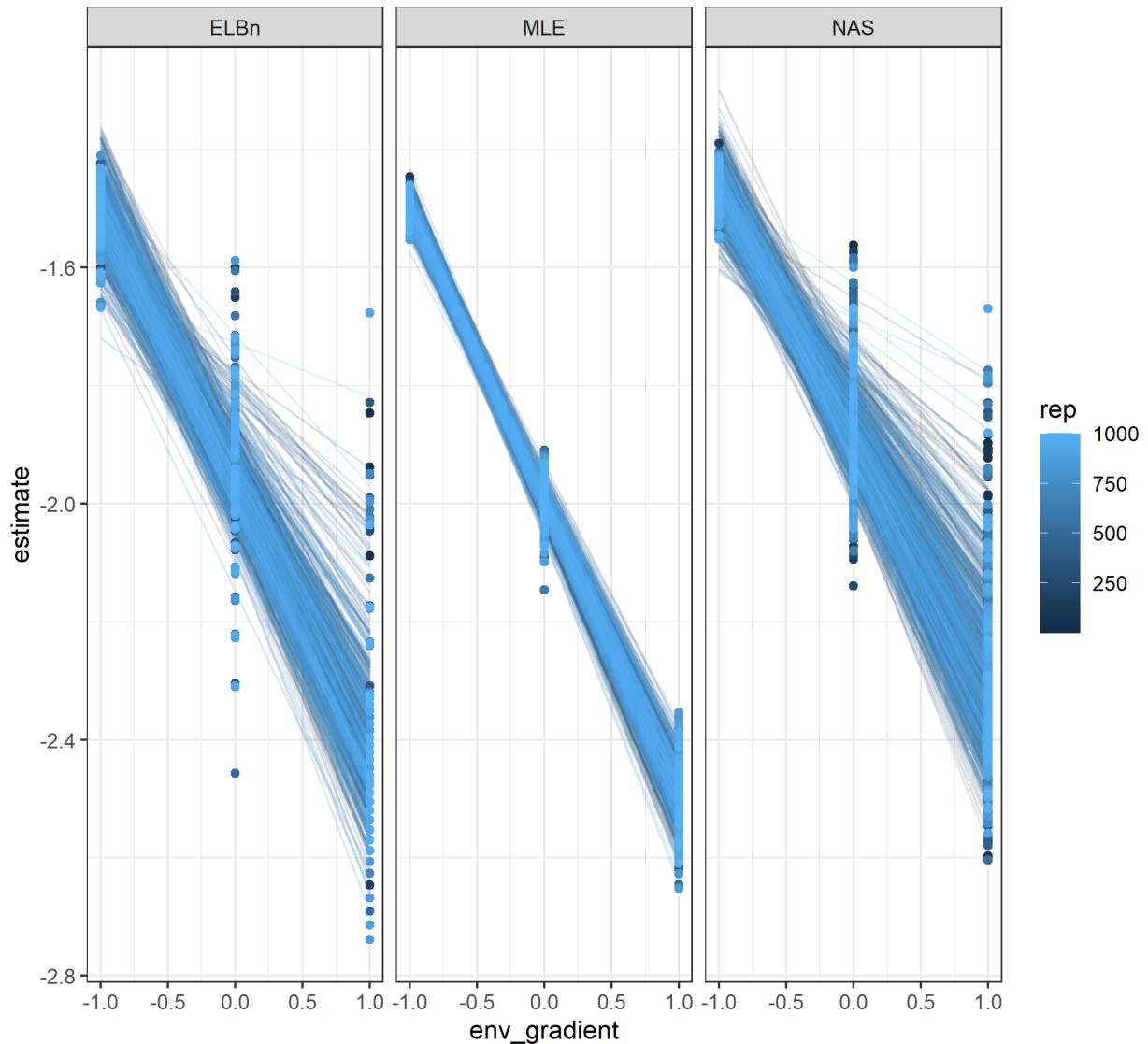


Figure S9: Individual regressions for three sites across a hypothetical gradient with a known relationship of 0.5. All other parameters are the same as in the main analysis.

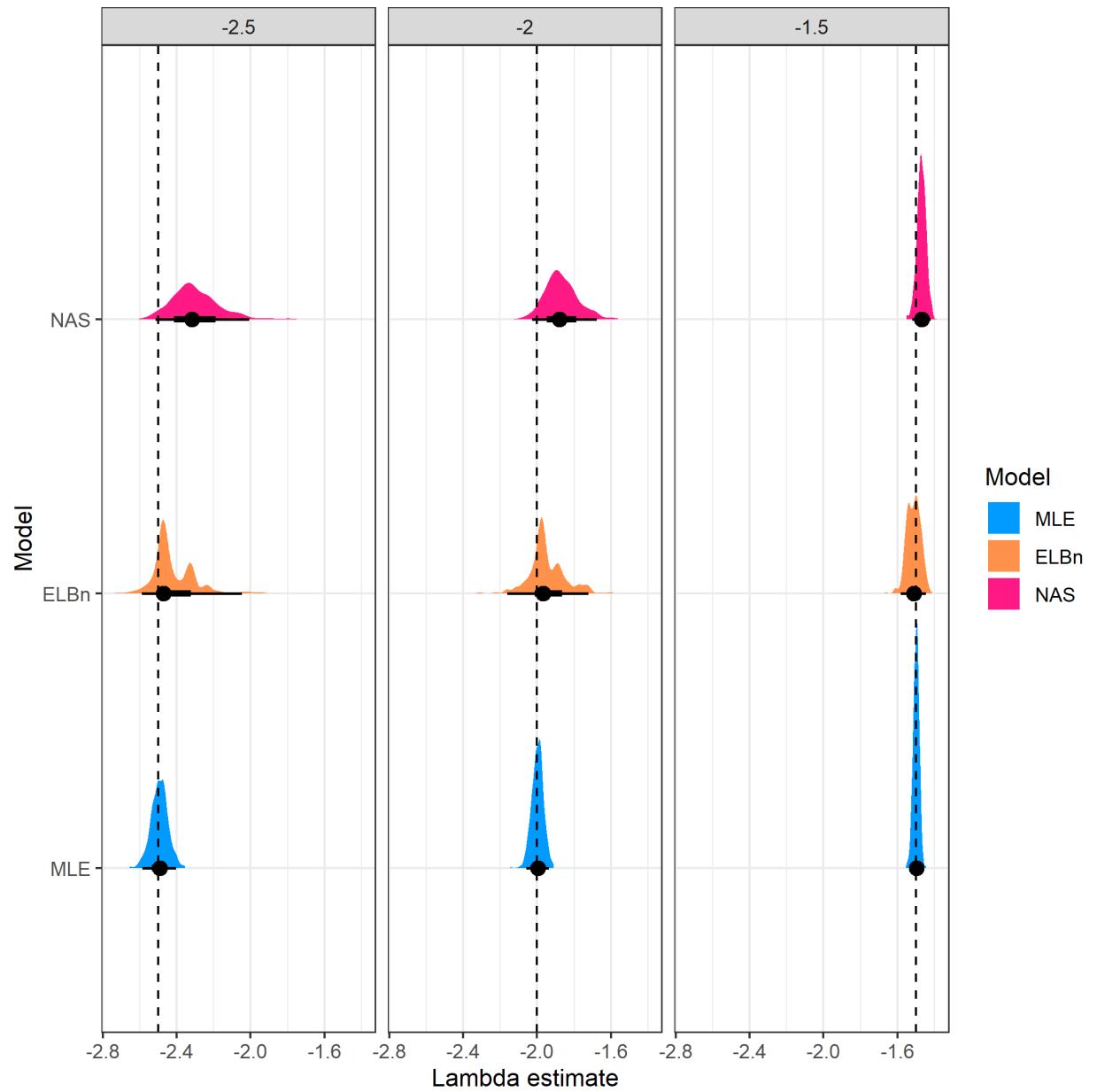


Figure S10: Distribution of estimated  $\lambda$  coefficient for three sites across a hypothetical gradient with known values. All other parameters are the same as in the main analysis.

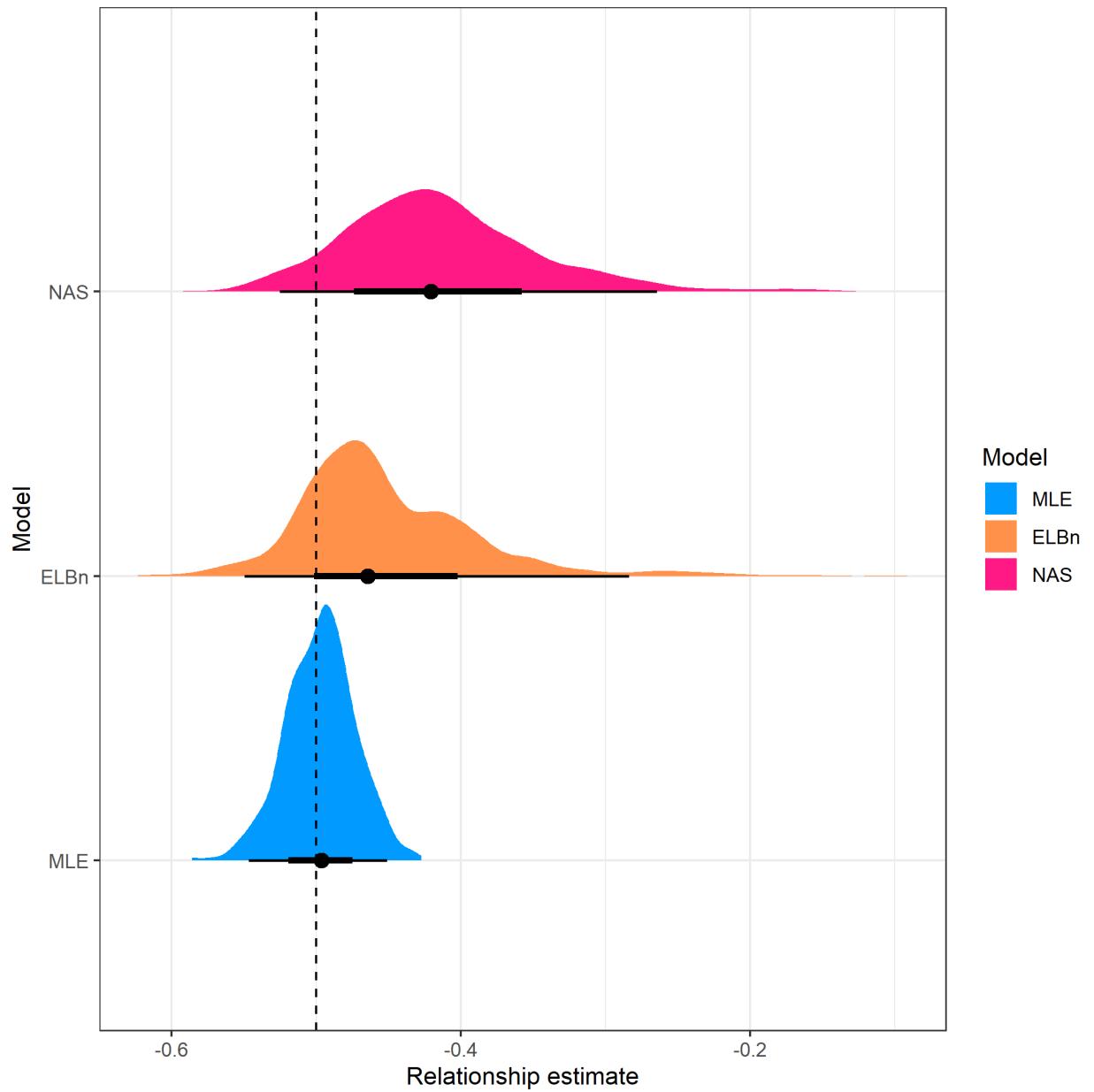


Figure S11: Distribution of estimated relationship ( $\beta_1$ ) coefficient's for three sites across a hypothetical gradient with known value of 0.5. All other parameters are the same as in the main analysis

<sub>312</sub> Sample size

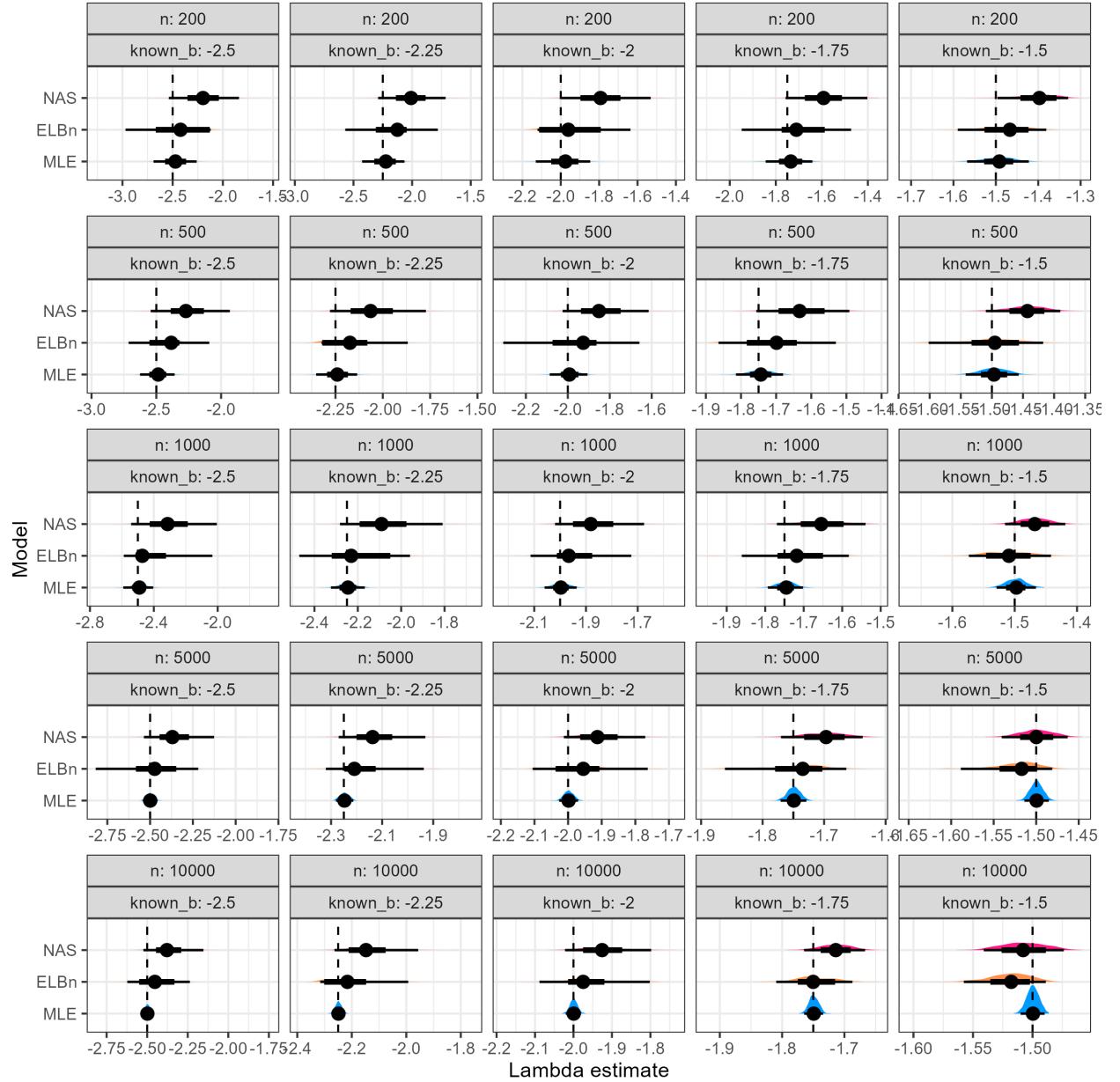


Figure S12: Distribution of size spectra parameter estimates. Vertical line is the known parameter (dashed line) which describes the bounded power law distribution from which the body size estimates were sampled. As  $n$  increases (top to bottom) and  $\lambda$  increases (left to right), the accuracy of the estimate improves across all methods.

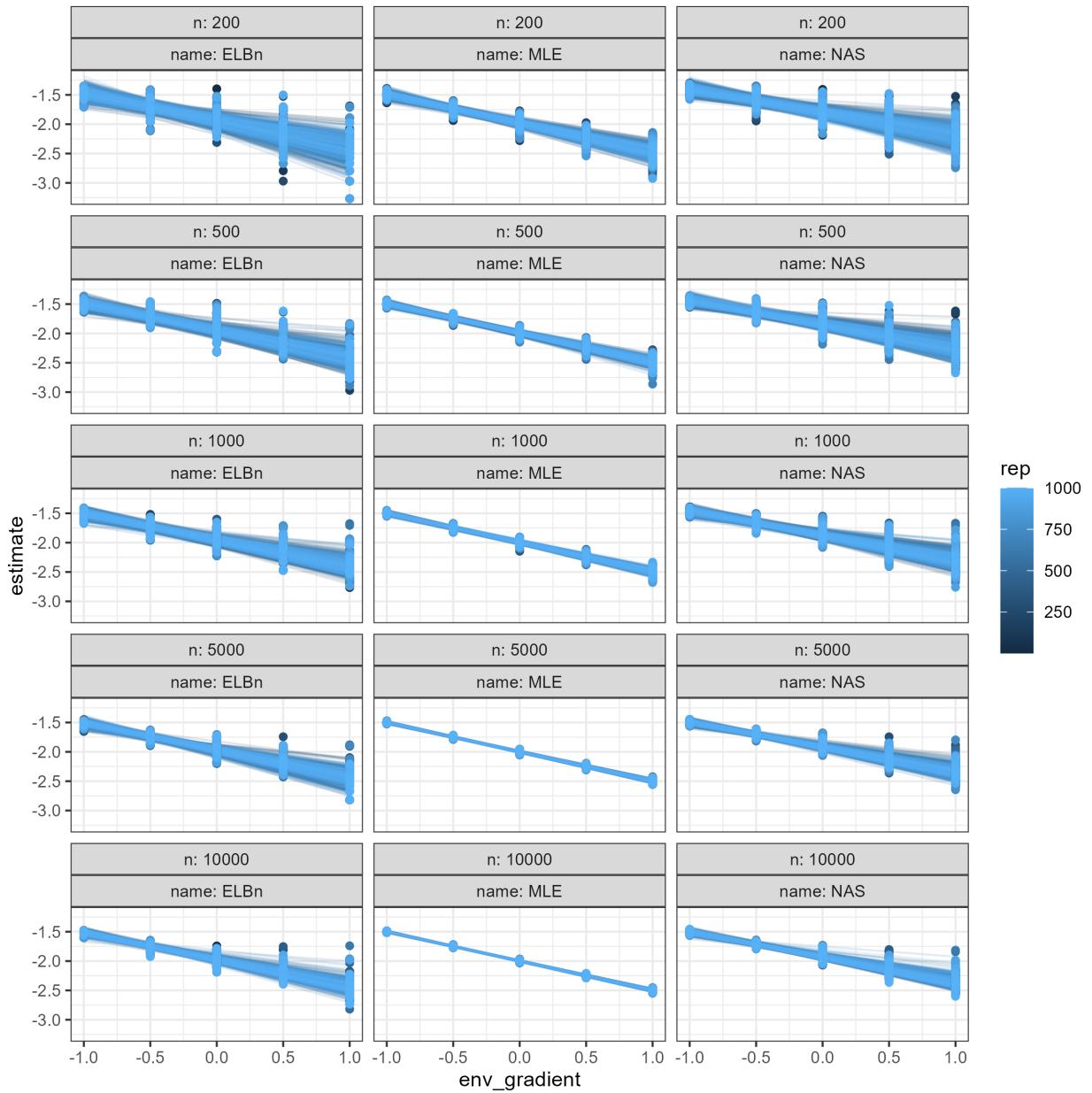


Figure S13: Individual regressions when varying sample size

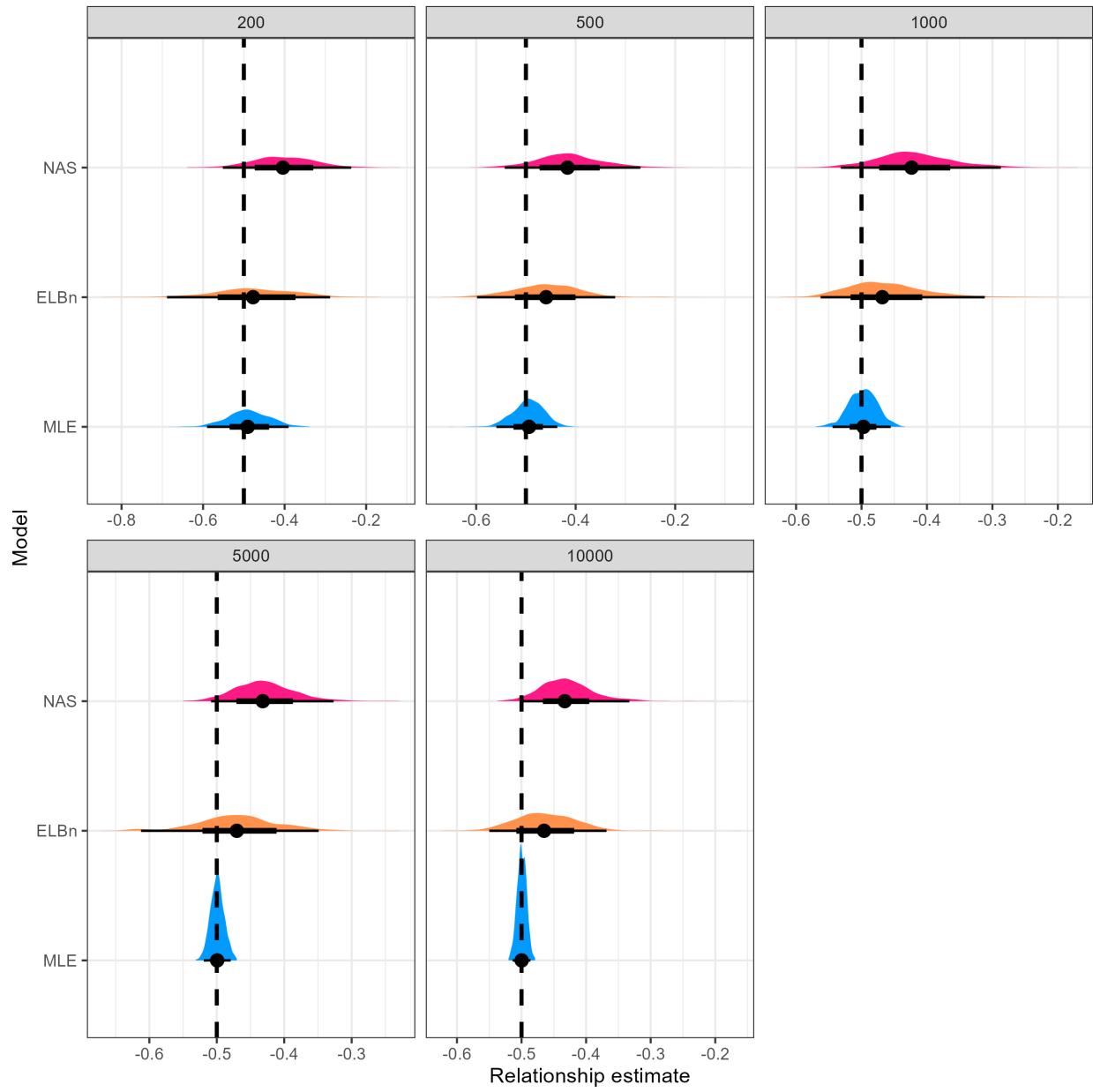


Figure S14: Distribution of the relationship coefficients with varying sample size

313  **$\lambda$  and relationship estimates**

314 In the main analysis, we simulate body size data from bounded power law distributions while varying  
315 the  $\lambda$  exponent which describes the distribution. For the results presented in the main text, we held the  
316 number of sites at 5, scaled the environmental gradient from -1 to 1, and set the minimum and maximum  
317 body sizes to 0.0026 and  $1.2 * 10^3$  respectively. Here, we plot the results of varying the number of sites (3,  
318 10), increasing the scale of the environmental gradient (-100 to 100) and decreasing the range of body sizes  
319 (min = 1, max = 100). Generally, the results reported in the main manuscript are robust to changing these  
320 parameters: the MLE estimate is nearly always closer to the known parameters, and the variation in these  
321 estimates is usually smaller than the binning methods.

322 Number of sites = 3

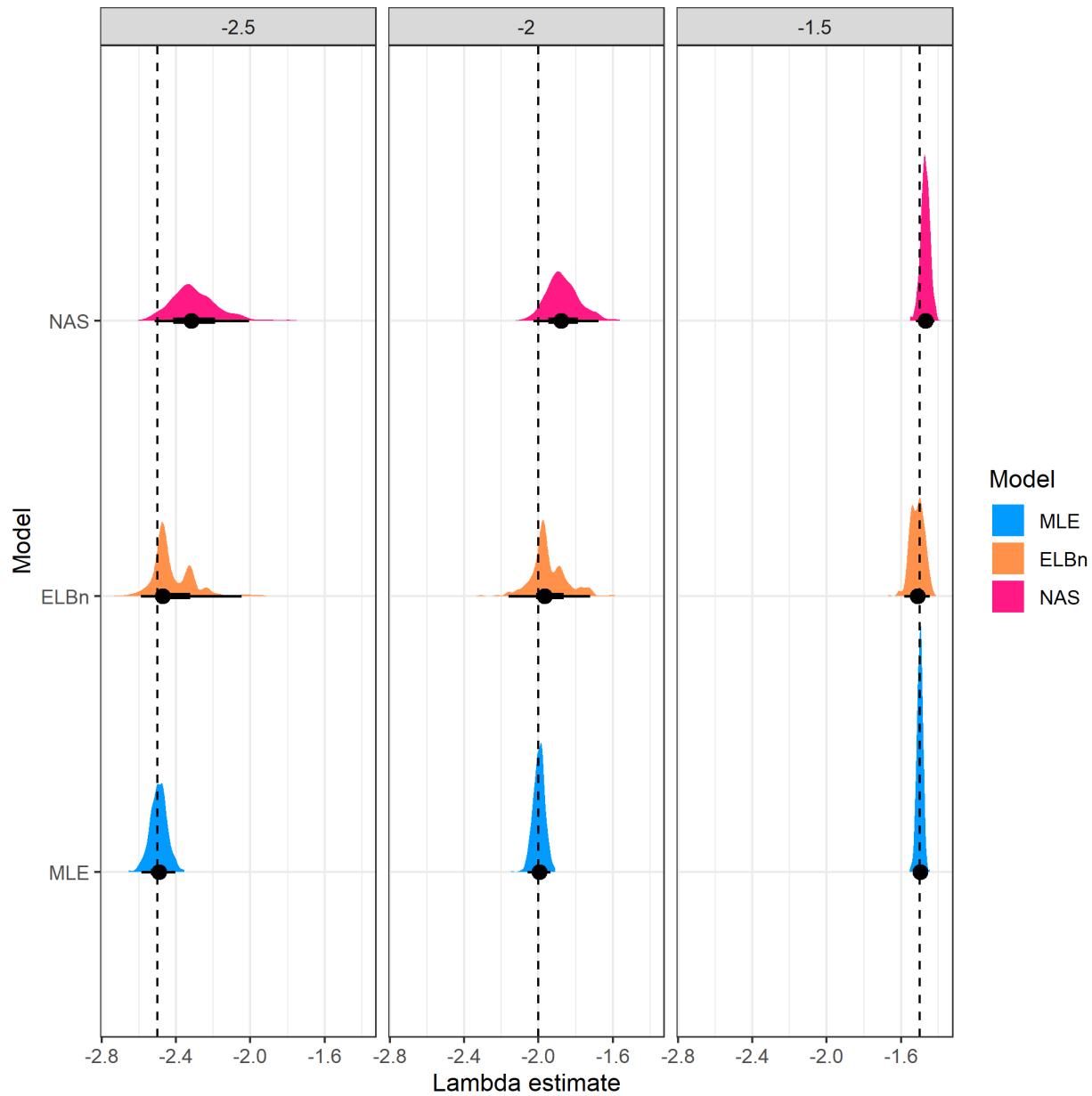


Figure S15: Distribution of estimated  $\lambda$  coefficient across a hypothetical gradient with known values (dashed line). The number of sites is reduced to 3, compared with 5 sites presented in the main analysis.

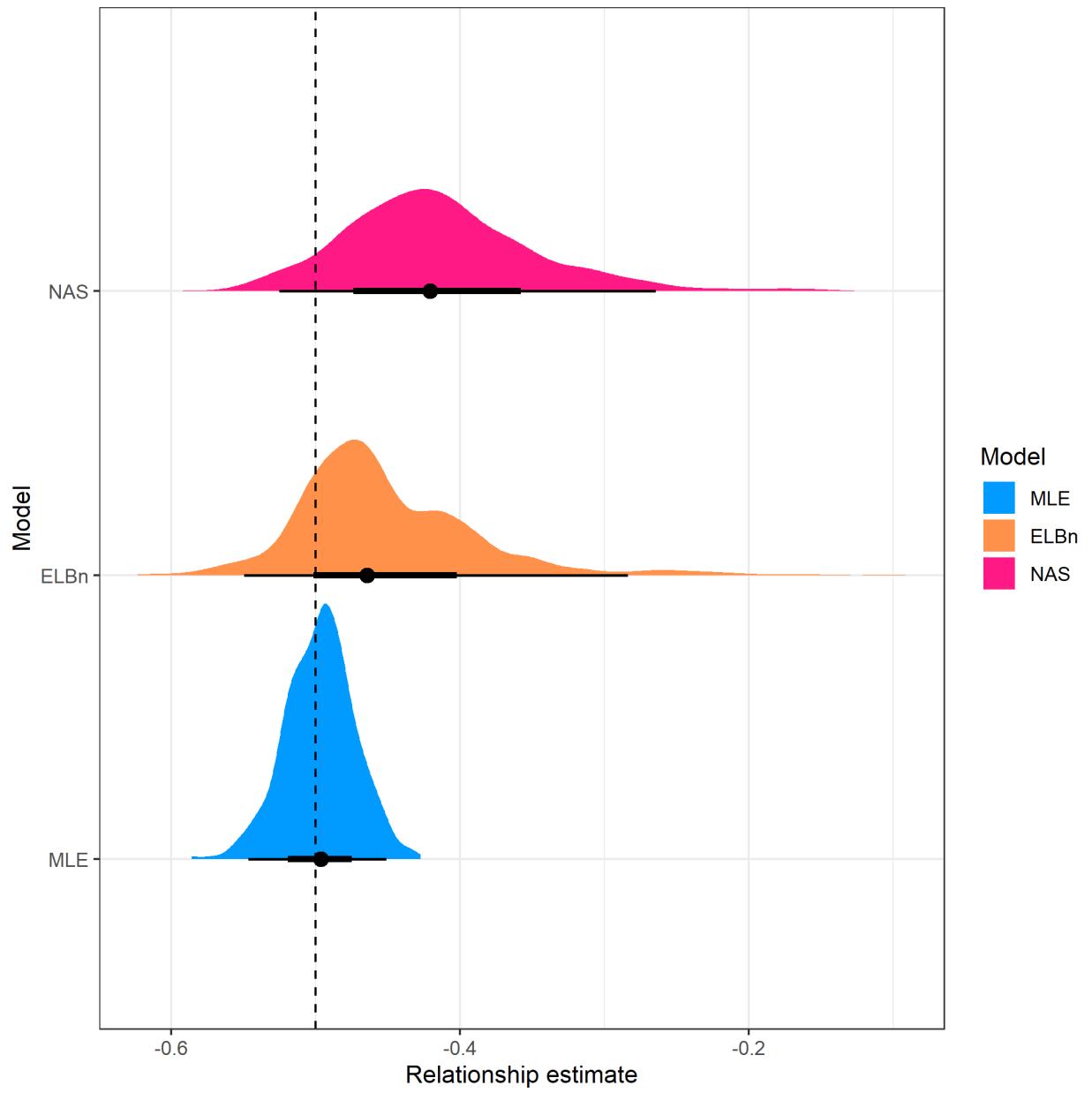


Figure S16: Relationship estimate

323 Number of sites = 10

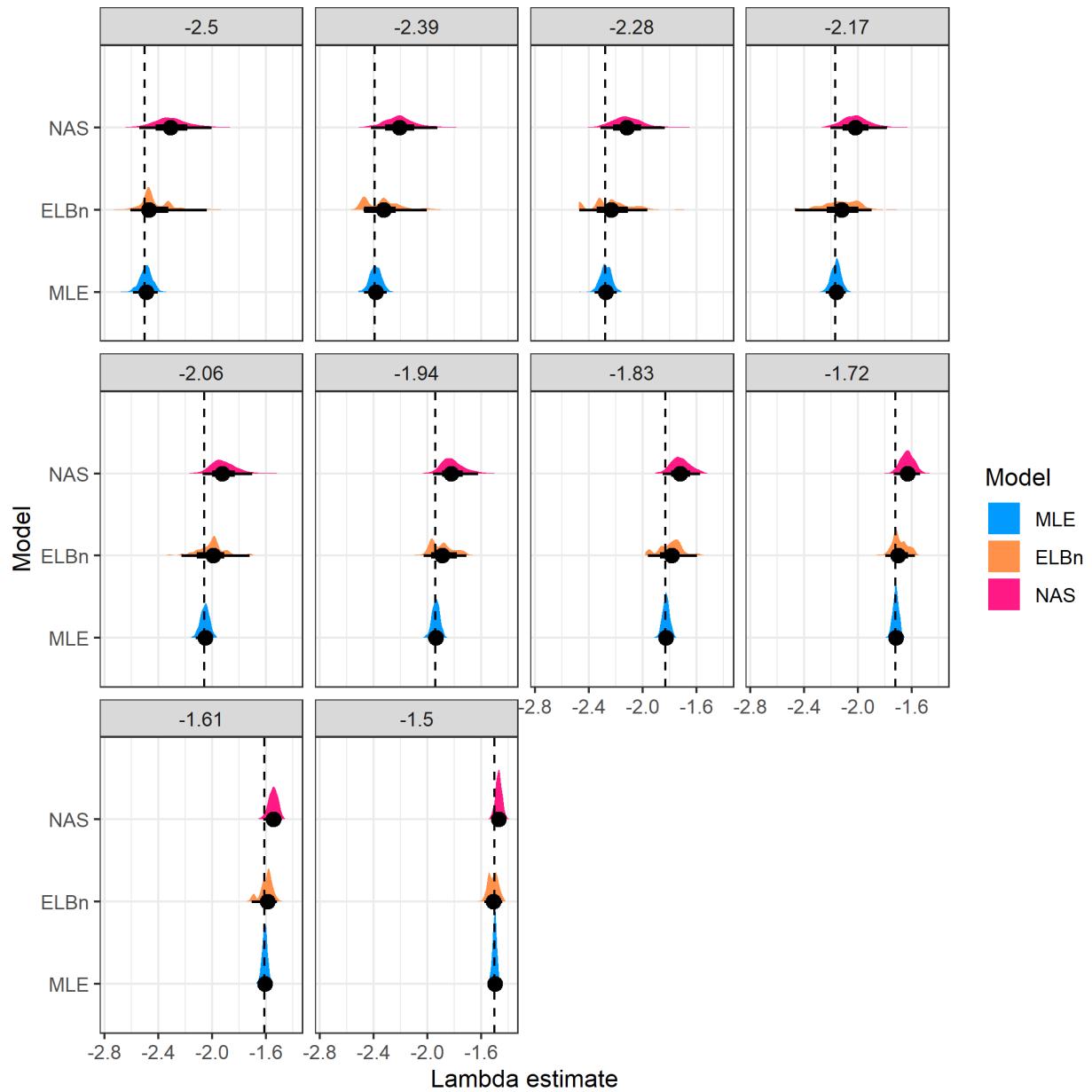


Figure S17: Distribution of estimated  $\lambda$  coefficient across a hypothetical gradient with known values(dashed line). The number of sites is increased to 10, compared with 5 sites presented in the main analysis.

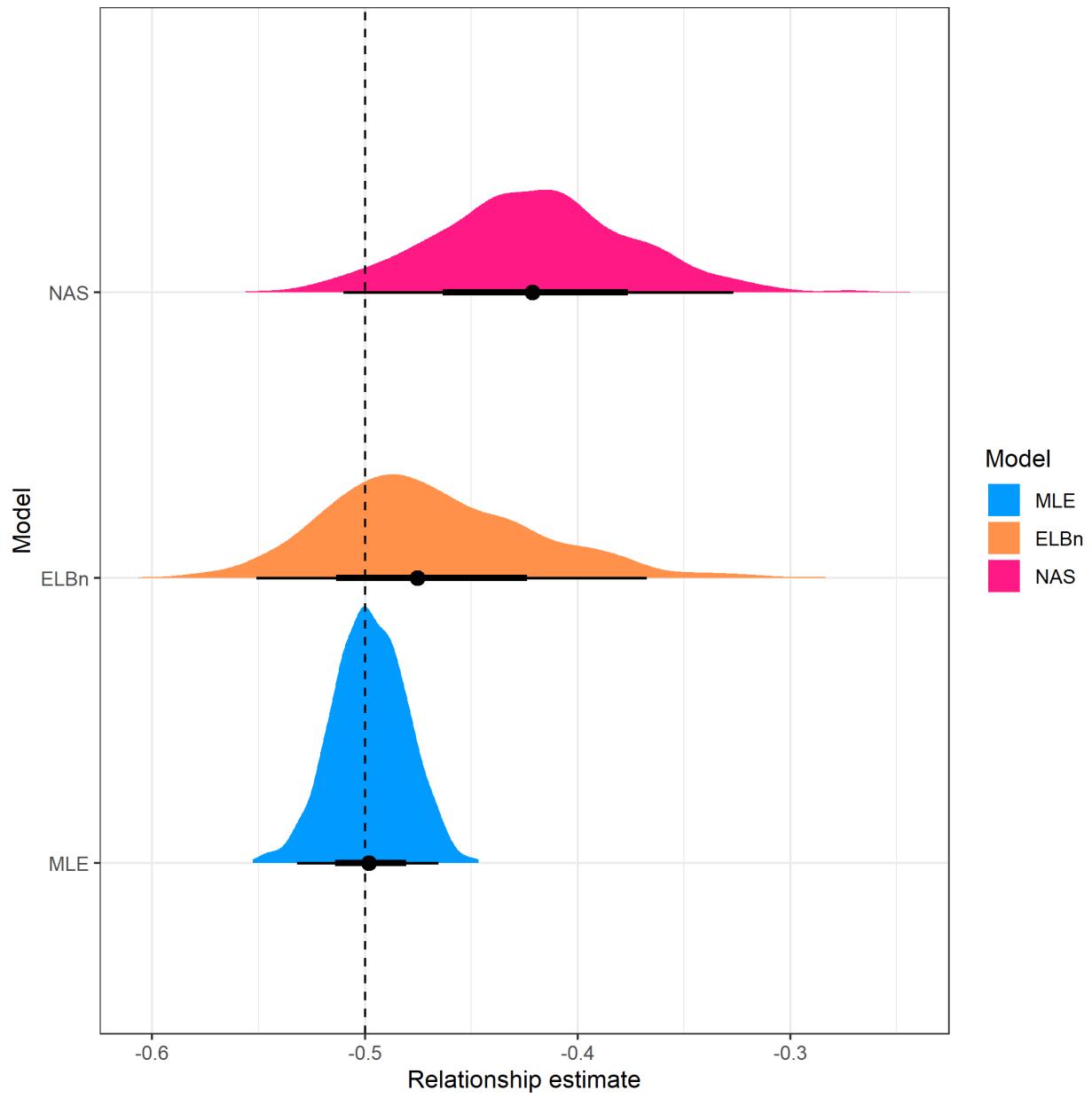


Figure S18: Relationship estimate

324 Environmental Gradient

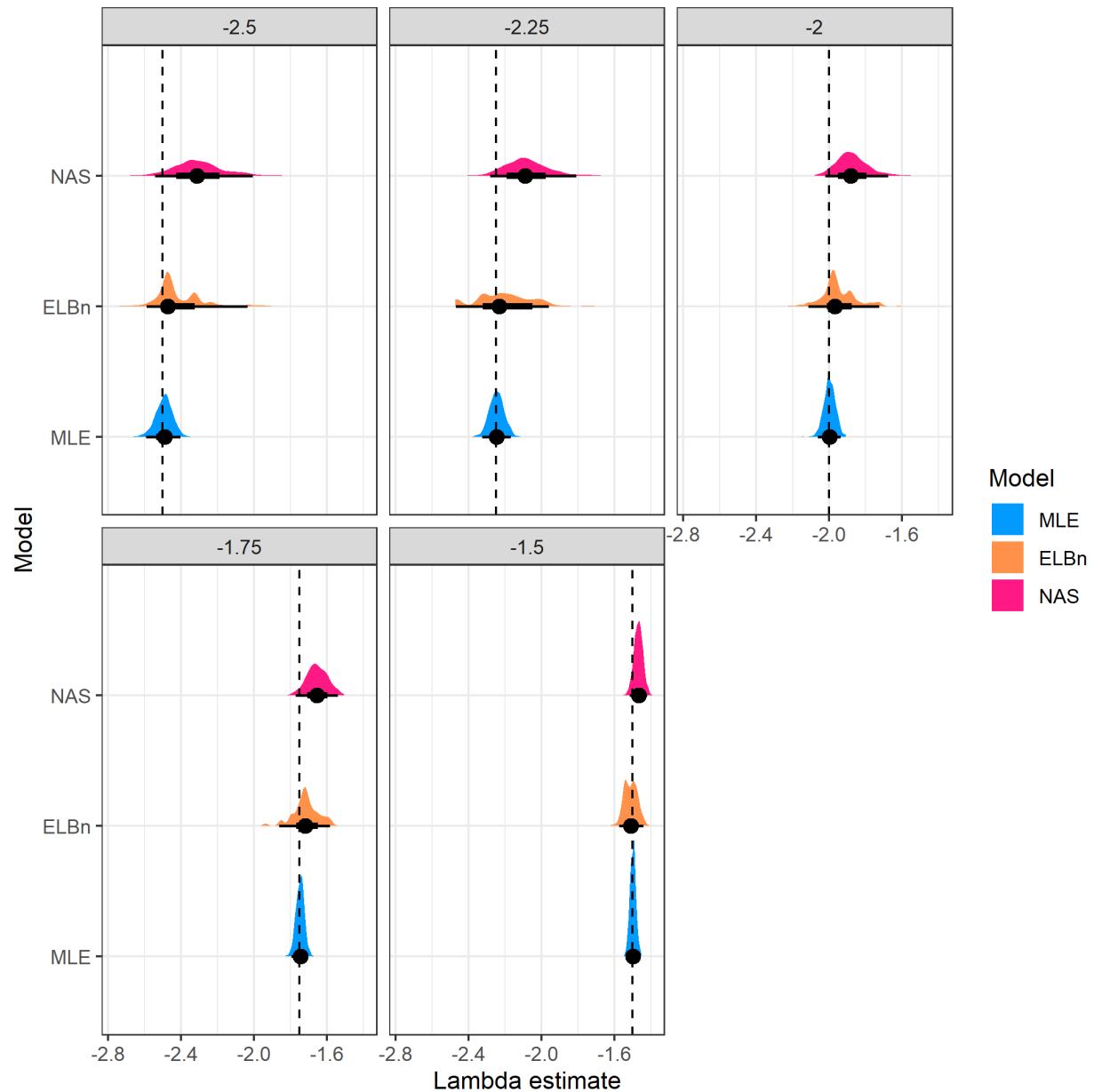


Figure S19: Distribution of estimated  $\lambda$  coefficient for five sites across a hypothetical gradient with known values (dashed line). The range of the environmental gradient was increased (-100 to 100).

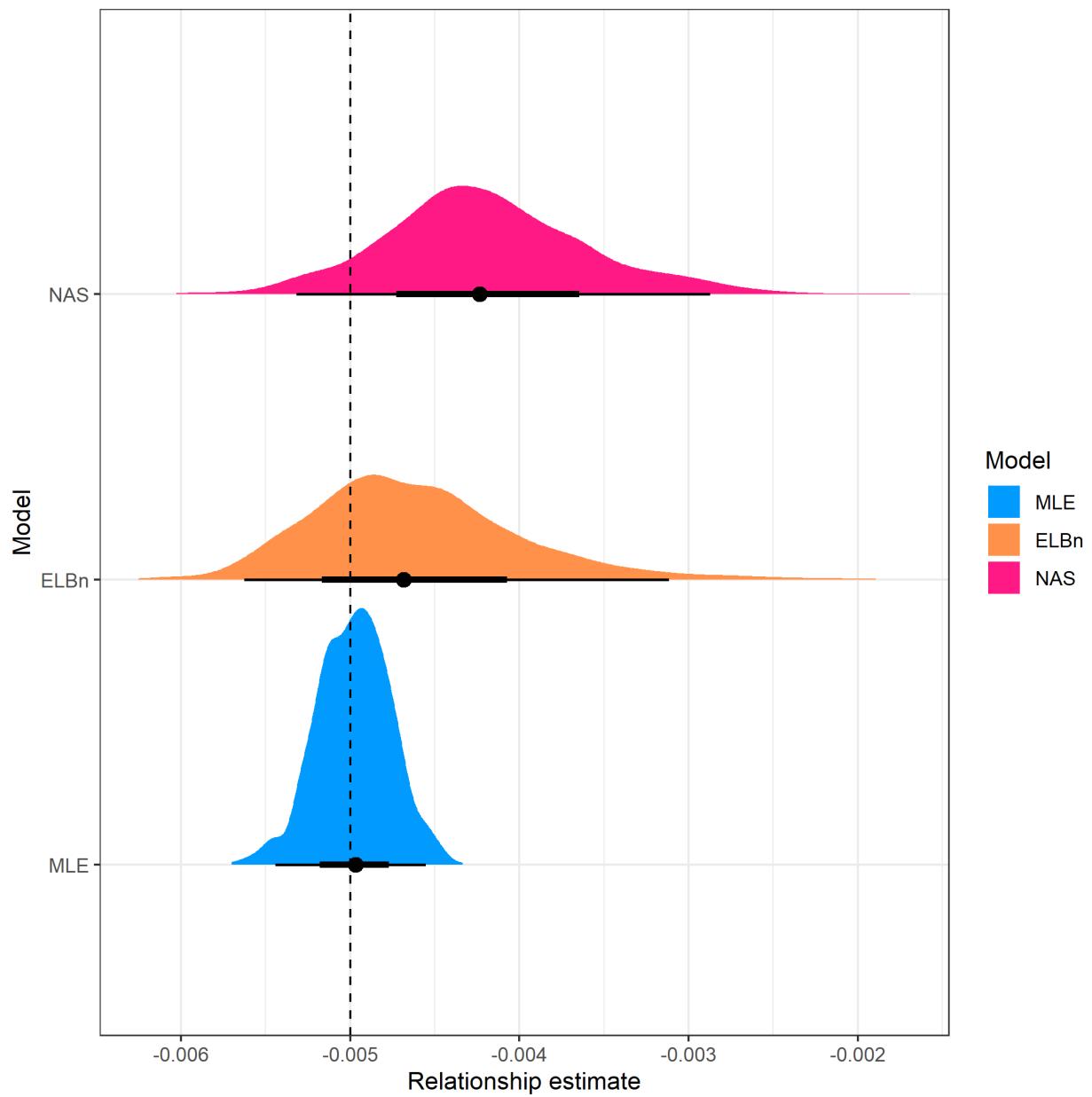


Figure S20: Relationship estimate

325 Range of body sizes = 1 to 100

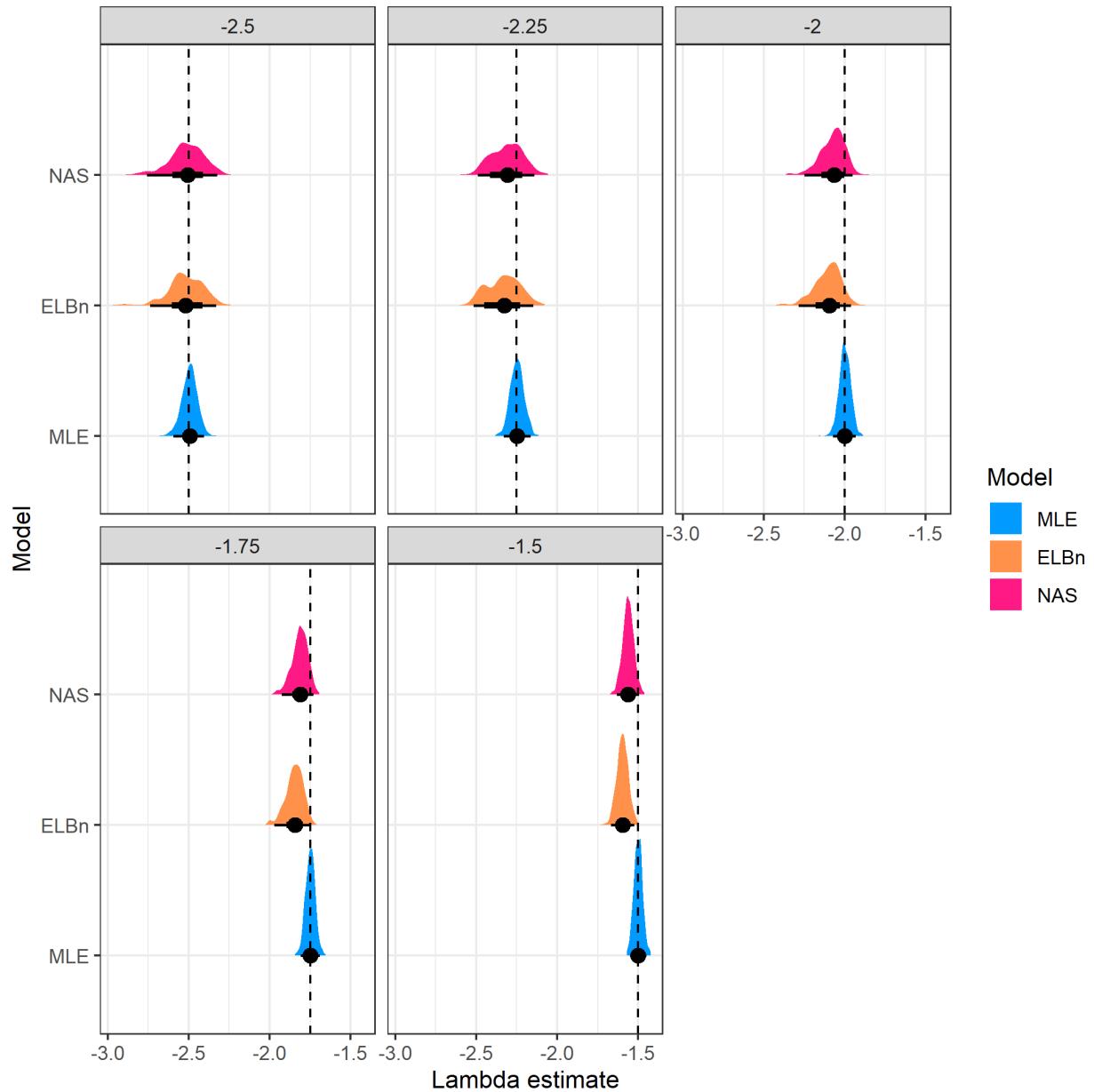


Figure S21: Distribution of estimated  $\lambda$  coefficient for five sites across a hypothetical gradient with known values(dashed line). Range of body sizes is smaller than main analysis and ranges from 1, to 100.

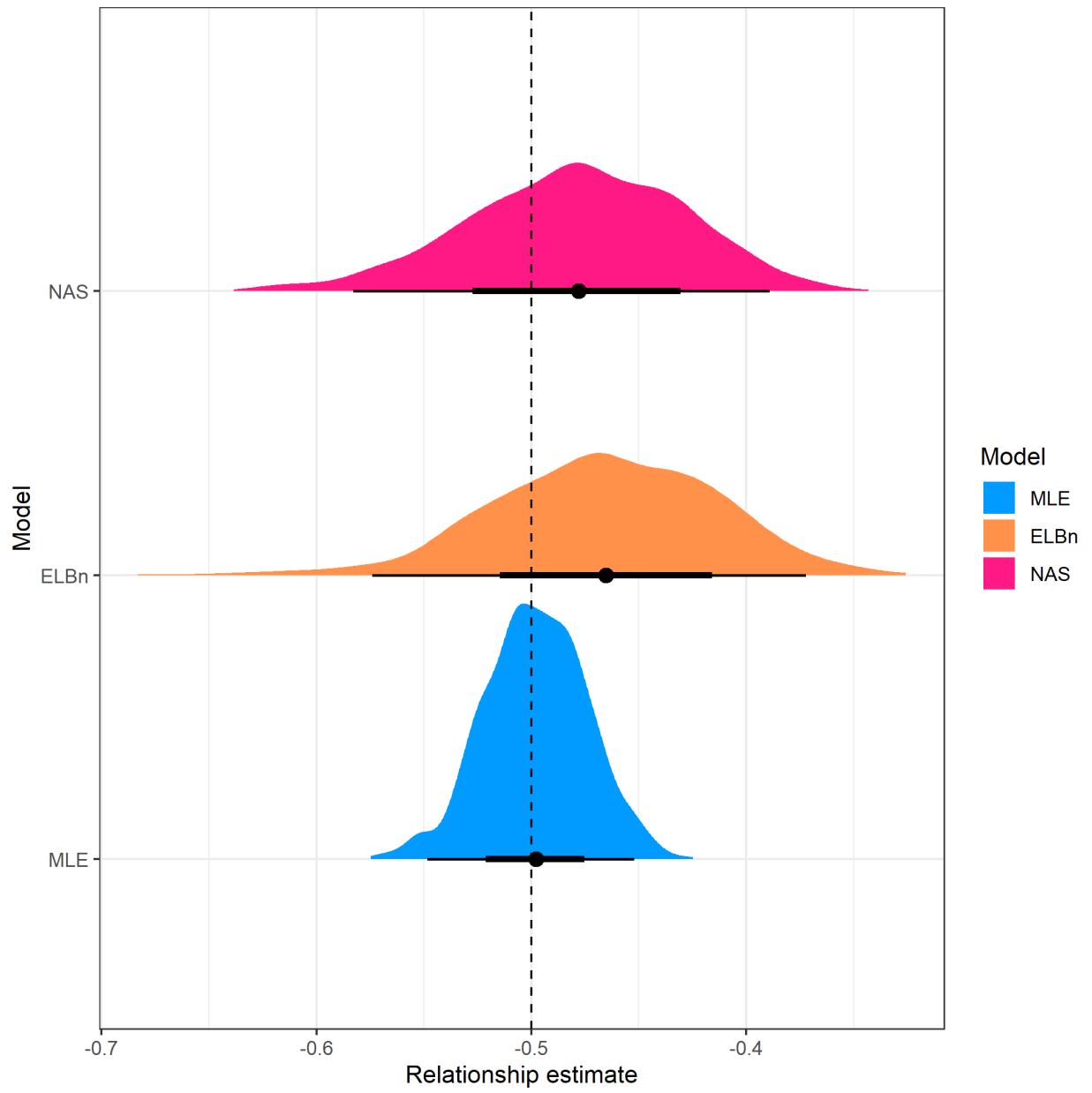


Figure S22: Relationship estimate