

Manuscript: Detecting differences in Size Spectra

Justin Pomeranz^{1,*} James R. Junker^{2,3} Vojisava Gjoni⁴ Jeff S. Wesner⁴

24 September, 2022

Contents

5	Abstract	2
6	Introduction	2
7	Methods	3
8	Data Simulation	3
9	Finer resolution of λ values	3
10	Estimation of Size spectra parameter λ	4
11	Estimation of relationship across gradient, β_1	4
12	Empirical Data	4
13	Data Availability	4
14	Results	7
15	Relationship across the hypothetical environmental gradient	7
16	No relationship	11
17	Small variation in lambda	11
18	Empirical data	11
19	Simulating shallower lambdas	15
20	Discussion	17
21	Concluding Remarks	18
22	References	19
23	Supplementary material	20
24	Range of body sizes, M	20
25	Large environmental gradient	22
26	Varying number of sites	25
27	Sample size	31
28	λ and relationship estimates	35
29	Comparison with other published estimates	43

¹ Colorado Mesa University, Grand Junction, CO, USA

² Great Lakes Research Center, Michigan Technological University, Houghton, MI USA

³ Louisiana Universities Marine Consortium, Chauvin, LA USA

⁴ Dept. of Biology, University of South Dakota, Vermillion, SD, USA

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

NOTE I'm still working on integrating with Zotero and importing all the necessary references (thanks

Jim for the help and instructions!). Don't worry about the ?? references in the rendered version for now.

Savina I'm particularly keen to get your feedback on this. I think you mentioned including some other references and discussion in Jeff's Bayesian model paper. If you think they fit in here please let me know or

39 feel free to include them. Likewise, please include any other comments you have, or if you have any insight
40 into how the interpretation of results from different methods would vary.

41 Abstract

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

61 Introduction

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

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

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

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

97 Methods

98 Data Simulation

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

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

102 $M_{min} = 0.0026$ and $M_{max} = 1.2 * 10^3$. These values are based on empirical body sizes of stream benthic
103 communities reported in Pomeranz et al. (2020)(Pomeranz et al. 2022). Our results are not dependent on
104 the range of body sizes, and we show the results of other body size ranges in the supplemental information.

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

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

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

111 Sample size, n

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

114 “No” relationship

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

121 Finer resolution of λ values

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

129 **Estimation of Size spectra parameter λ**

130 For each sample of M and each of the 1000 replicates, we estimated the exponent λ using MLE methods
131 modified from the the `sizeSpectra` package (Edwards et al. 2020). In addition, we use two common binning
132 methods to estimate the OLS slope parameter in log-log space. For the first binning method, we created 6
133 equal logarithmic bins covering the range of body sizes. The count in each bin was normalized by dividing
134 by the bin width. This method has been used by (refs here, I think a lot of Woodward's group uses this...)
135 Throughout the manuscript, the normalized equal logarithmic binning method will be referred to as ELBn.
136 The second method was similar to ELBn, but bins of Log_2 widths are used. The count in each bin is
137 normalized in the same way. This method has been used by (Pomeranz et al. 2019, McGarvey et al. 2019)
138 and is referred to as the Normalized Abundance Spectrum (NAS). Previous work has investigated the biases
139 in λ estimates when using different methods (Edwards et al. (2017)(2017), Edwards et al. 2020(2020), White
140 et al. 2008(2007)). The focus of the present work is to investigate biases when estimating the relationship or
141 relative change in parameters across an environmental gradient. Hence, we do not discuss the λ estimates in
142 the main text, but the results are presented in the Supplemental Information (**NOTE** I'm thinking about
143 dropping this from SI. Might be too much info, and it's already been published on by Edwards. Does this
144 add anything to the story??).

145 **Estimation of relationship across gradient, β_1**

146 For each simulation replicate, we estimated how the size spectra parameters varied across the hypothetical
147 gradient. Simple OLS regression were conducted in the form $\lambda_{\text{estimate}} \beta_0 + \beta_1 * X$. The distribution of the
148 relationship coefficient, β_1 , were plotted compared to the known relationship.

149 **Empirical Data**

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

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

167 **Data Availability**

168 The data used in this manuscript is already publicly available (Pomeranz et al. 2019 data dryad DOI
169 [here](#), NEON data citation [here](#). R Scripts to reproduce the full simulation and analysis are available at:
170 ([<https://github.com/Jpomz/detecting-spectra-differences>]).

PLB

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

$b = (-2.5, -1.5)$

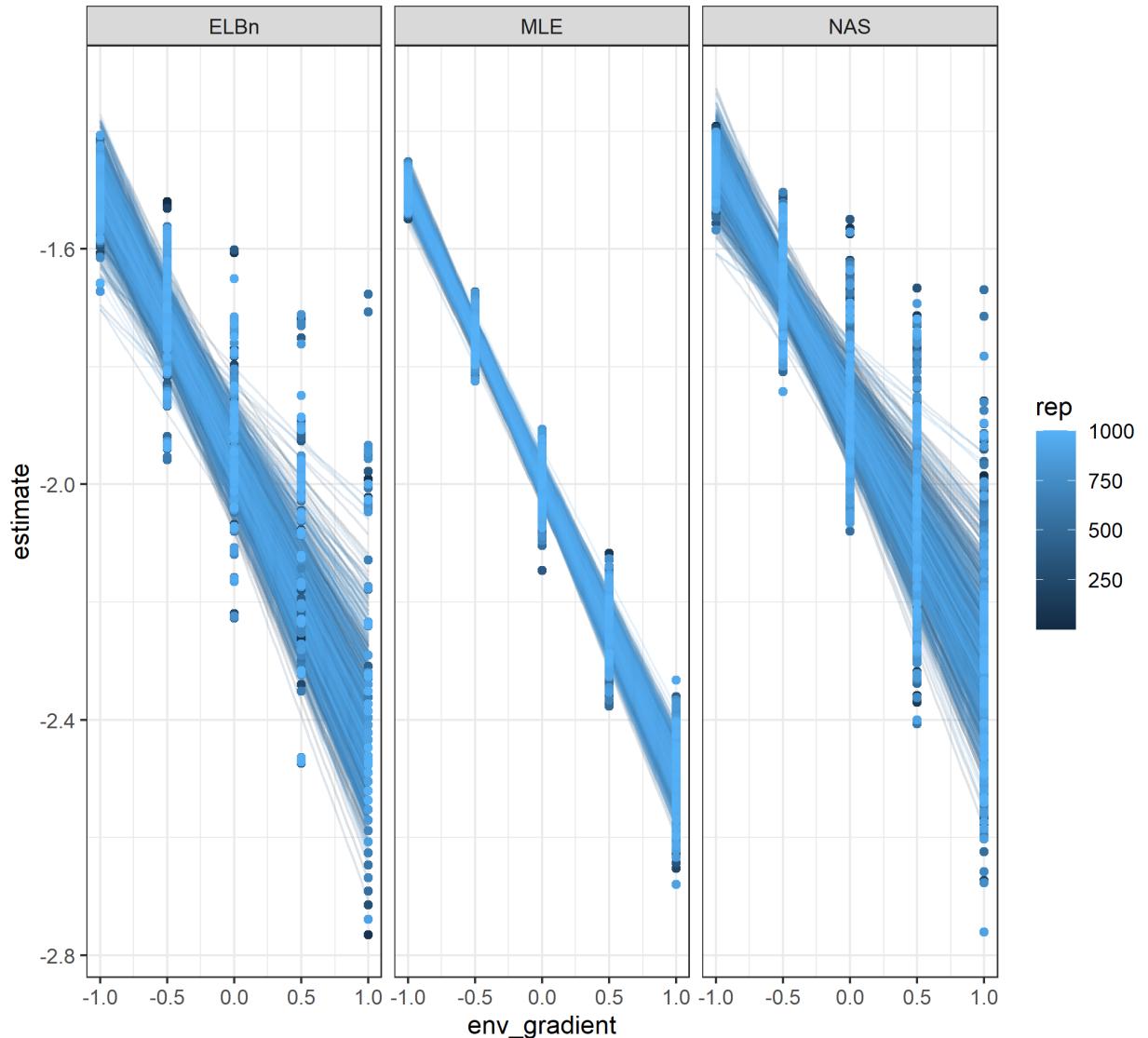


Figure 1: Relationship estimates across the hypothetical gradient for each replicate. Each panel is a different method for estimating the size spectra parameter.

PLB

$m_range = (0.0026, 1200)$
 $b = (-2.5, -1.5)$

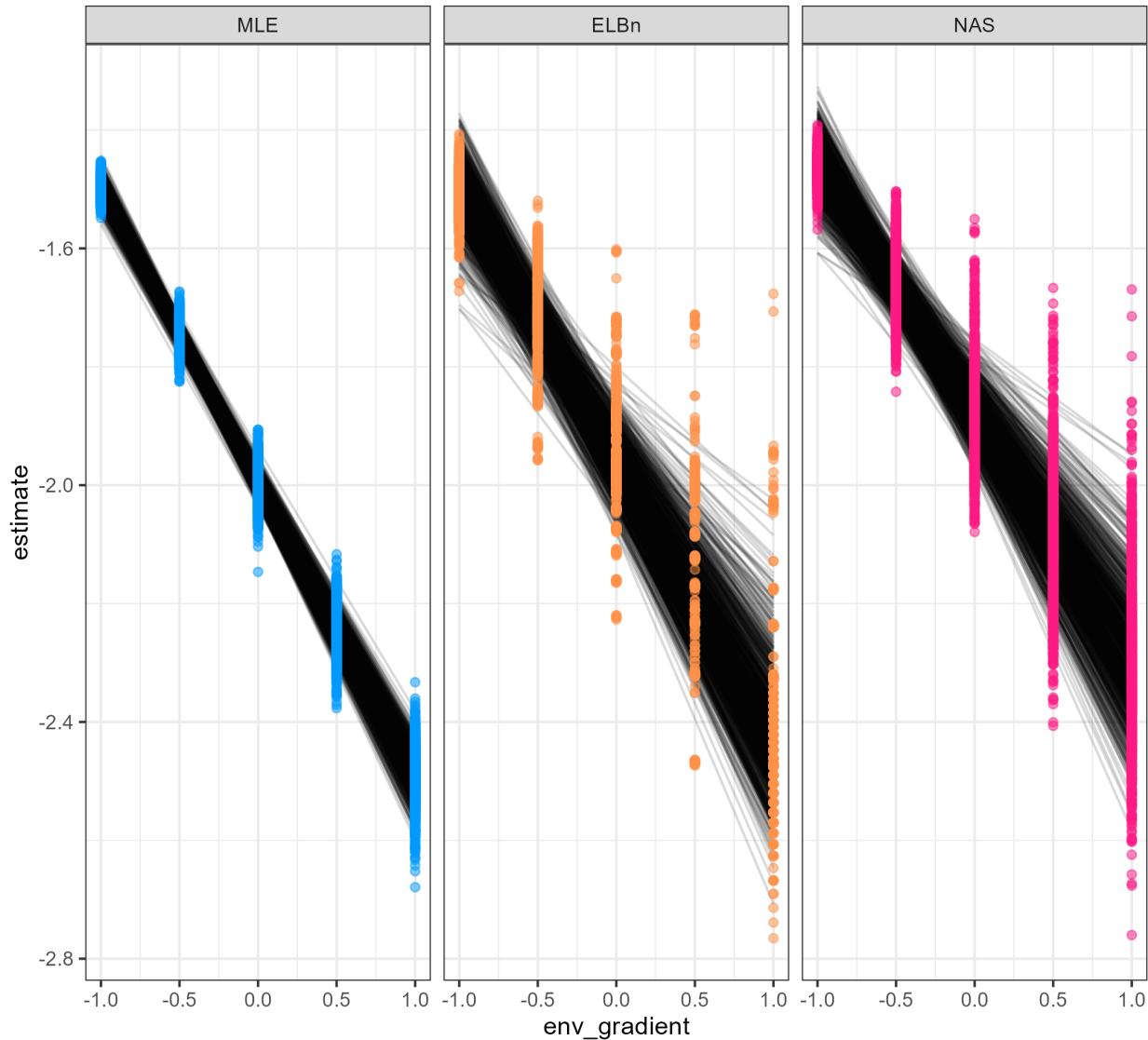


Figure 2: alternate plot version; I think I like this one better. What are your thoughts on which you prefer??

₁₇₁ **Results**

₁₇₂ **Relationship across the hypothetical environmental gradient**

₁₇₃ The distribution of β_1 estimates from MLE method were the most accurate and precise, and the NAS
₁₇₄ method was the least accurate and precise (Table *xx*). The median estimates for all methods were closer
₁₇₅ to the known relationship. All of the methods produced estimate distributions with long right tails, but
₁₇₆ this pattern was greater in the ELBn and NAS methods (e.g., SD for ELBn and NAS methods were 2-3x
₁₇₇ larger, Figure *xx* Halfeye plots). This result was consistent across simulation parameters, including sample
₁₇₈ size, number of sites, and λ values.

```
179 ##   Model    mean     p50     p25     p975
180 ## 1   MLE -0.495 -0.494 -0.499 -0.483
181 ## 2   ELBn -0.460 -0.489 -0.524 -0.291
182 ## 3   NAS -0.419 -0.442 -0.458 -0.302
```

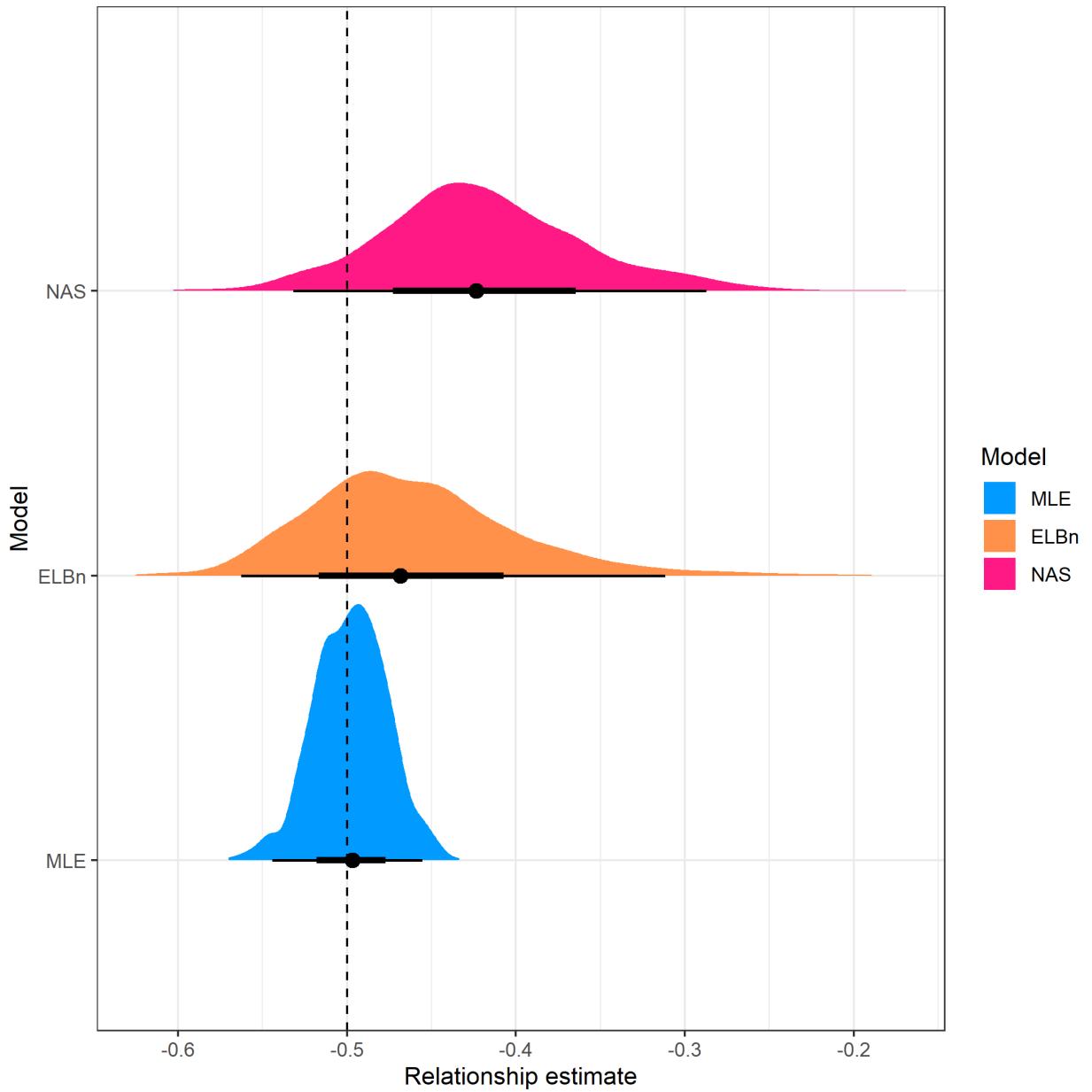


Figure 3: Halfeye plots for the estimated relationship coefficient estimates across 1000 replicates and three different methods. The point below each distribution plot is the median estimate, the thick and thin bars represent the 66th and 95th percentile, respectively. Vertical line is the known relationship.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

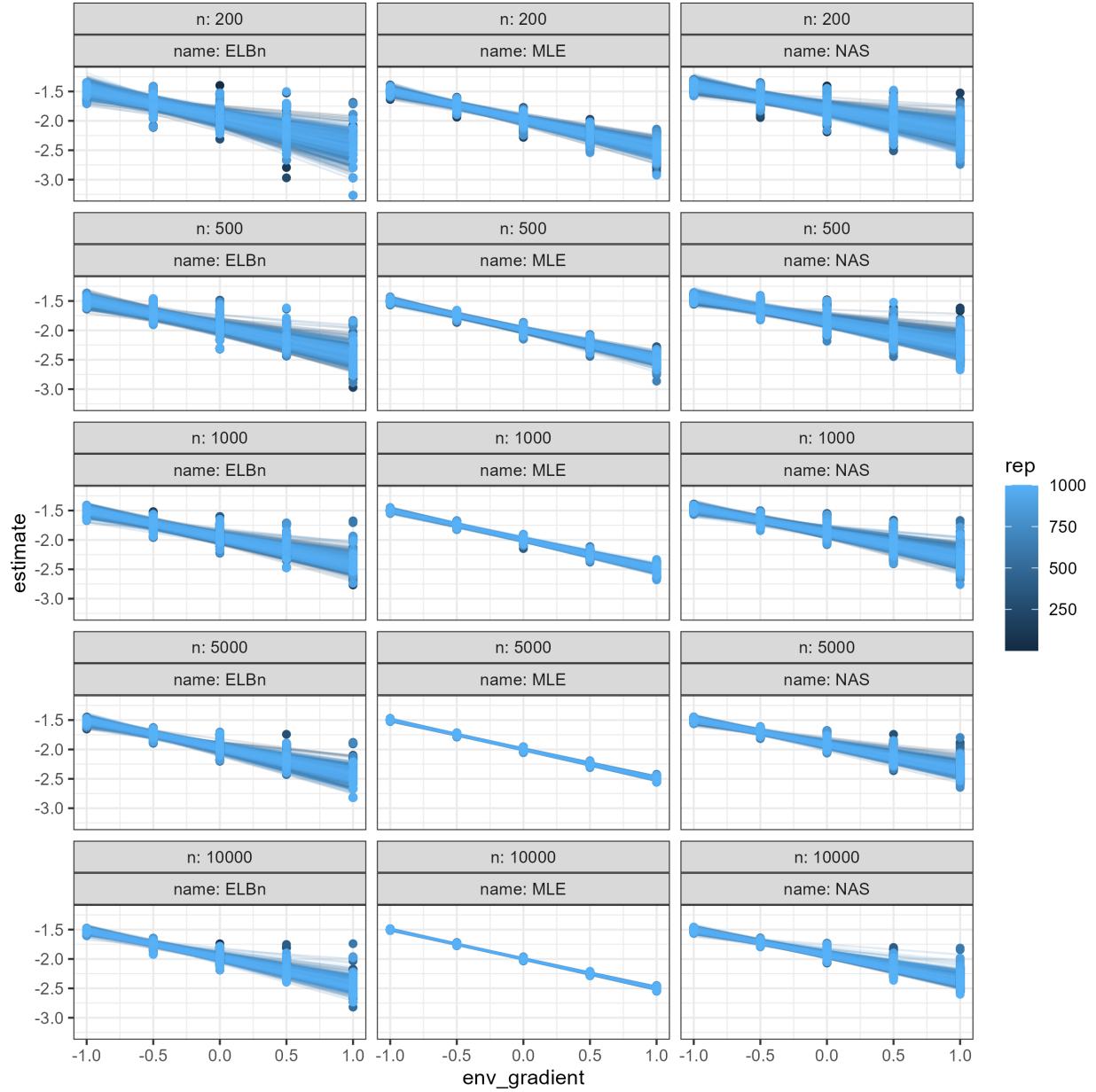


Figure 4: Individual regression estimates across the hypothetical gradient based on sample size (rows) and methodology used (columns). (match this figure to “new” style if we like that better)

183 **No relationship**

184 All of the methods performed similarly when there was no relationship across the hypothetical gradient.
185 The mean β_1 estimates for all three methods ranged from -0.0005 to 0.0002. Although the distributions were
186 closely centered on 0, the distribution of coefficient estimates was much wider for the two binning methods;
187 the 95th percentile of the MLE distributions is narrower than the 66th percentiles for the NAS and ELBn
188 methods.

```
189 ## Model mean p50 p25 p975  
190 ## 1 MLE -0.001 -0.001 -0.012 0.038  
191 ## 2 NAS 0.000 -0.001 -0.039 0.119  
192 ## 3 ELBn 0.000 -0.001 -0.046 0.134
```

193 **Small variation in lambda**

194 The performance of all methods to detect small variation in the λ parameter (between -1.9 to -2.1) across
195 the environmental gradient was also investigated. The median β_1 coefficient estimates for the MLE method
196 was -0.1, which exactly matched the value for the hypothetical relationship, while the other two methods
197 both had median estimates of -0.09. The upper 97.5 percentile of the distribution was below 0 for the MLE
198 method, meaning that most of the estimates had the correct sign. In contrast, the upper 97.5 percentile for
199 both the NAS and ELBn methods were greater than 0 at 0.027 and 0.052, respectively. (Figure 4).

```
200 ## Model mean p50 p25 p975  
201 ## 1 MLE -0.100 -0.101 -0.112 -0.062  
202 ## 2 ELBn -0.091 -0.095 -0.134 0.052  
203 ## 3 NAS -0.088 -0.091 -0.126 0.027
```

204 **Empirical data**

205 For both empirical data sets, the direction and magnitude of change (i.e. β_1 coefficients) are generally in
206 agreement. Size spectra parameters consistently increase (become flatter) in the AMD data (Fig. XX A).
207 Likewise, the size spectra parameters consistently increase (become steeper) with increasing temperature
208 across the NEON sites (Fig XX D)

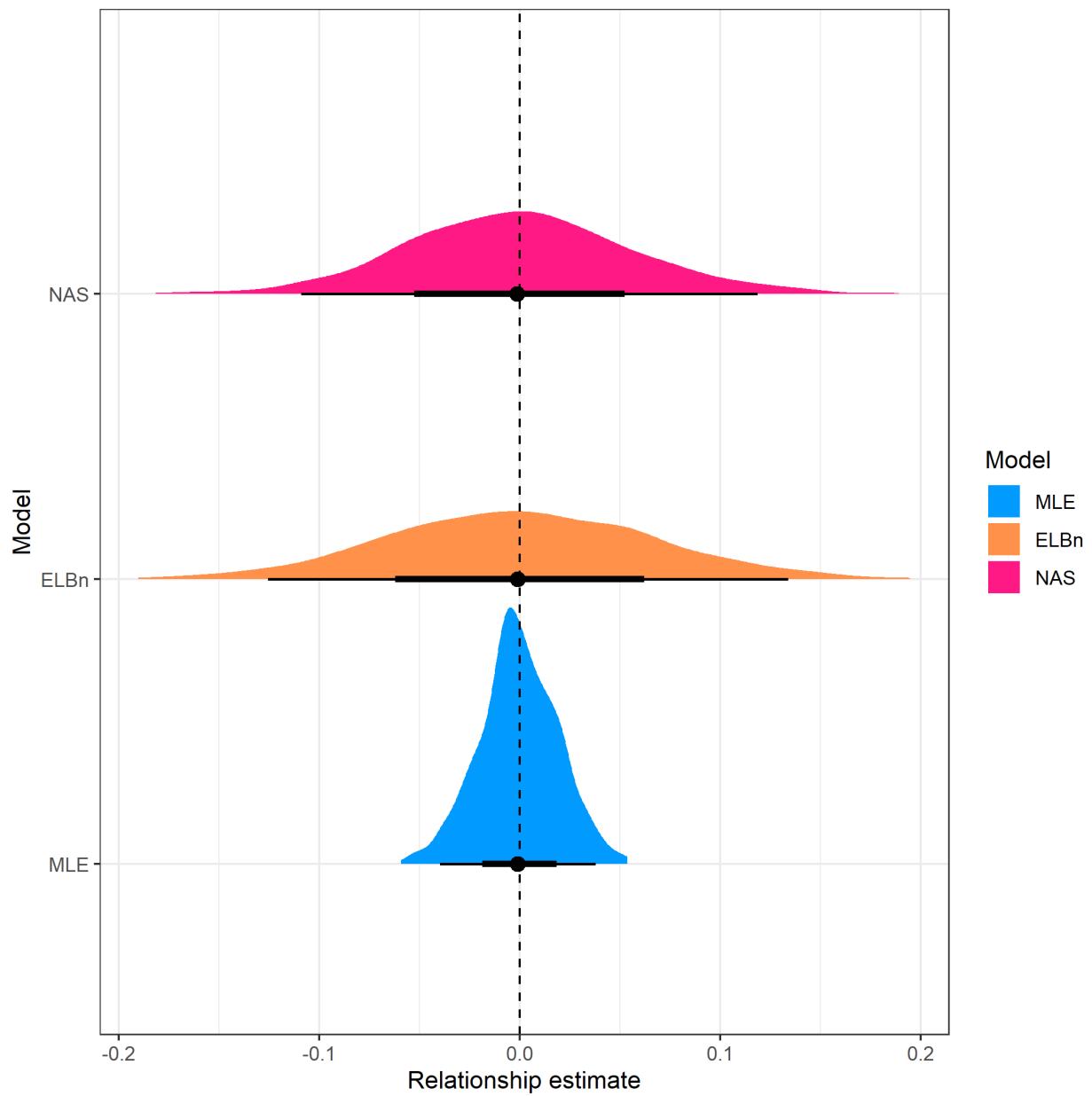


Figure 5: Individual regression estimates when no relationship exists across the hypothetical environmental gradient. The methods are separated by rows. The median for all methods is approximately zero, but the distribution for the MLE method is narrower than the two binning methods.

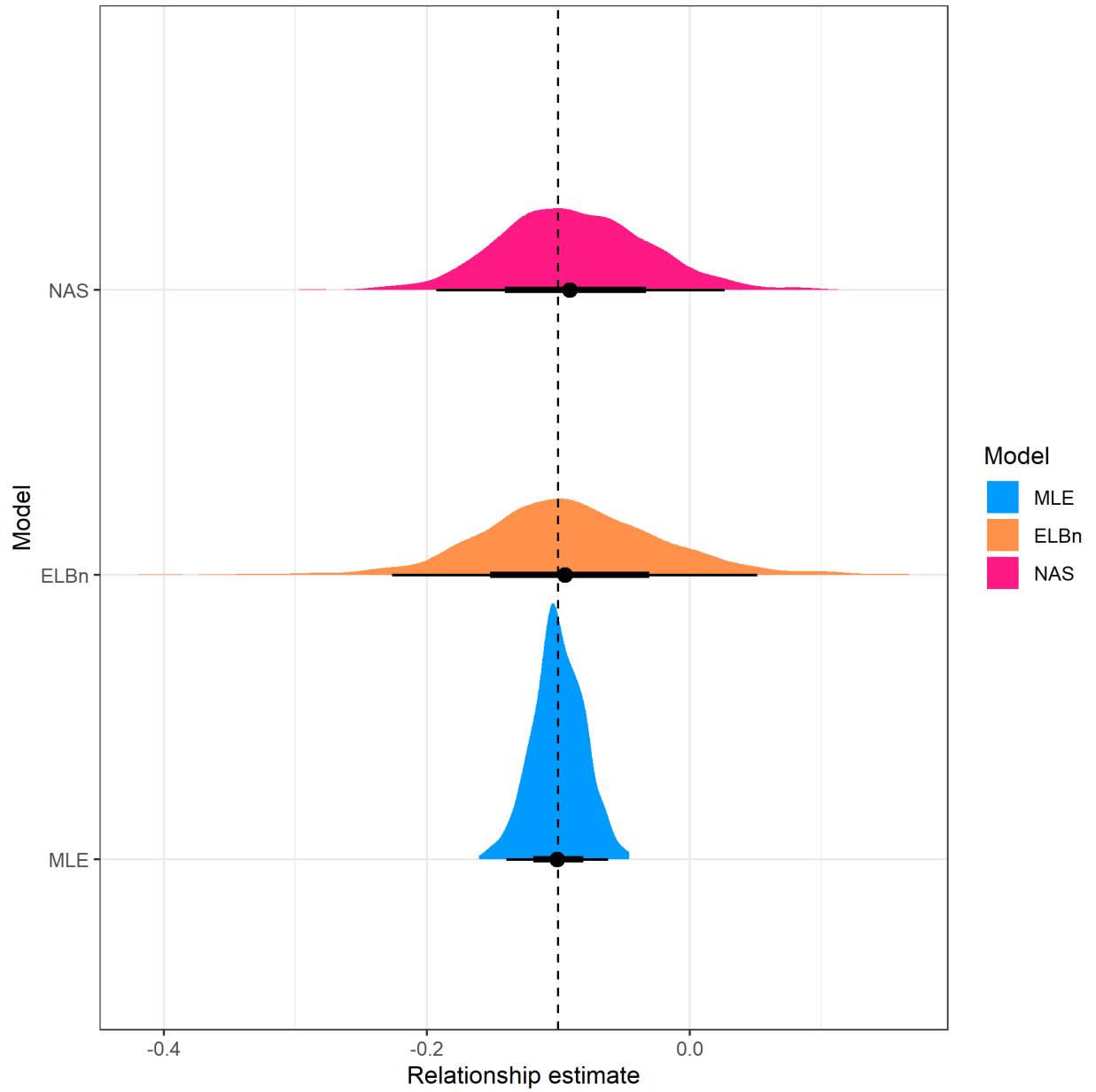


Figure 6: Distribution of the estimated relationship for each method. Here, λ varied from -1.9 to -2.1, with a known relationship of -0.1 across the gradient indicated by the dashed vertical line. The distribution of relationship estimates was approximately centered at the known relationship for all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

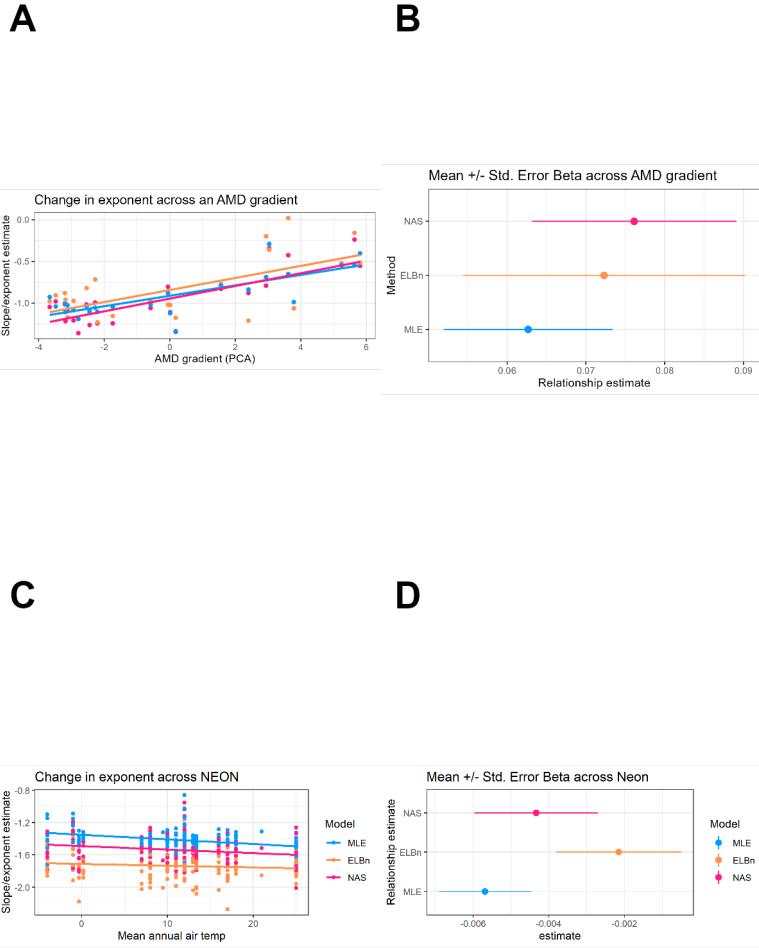


Figure 7: Figure of empirical data estimates. All of the methods estimate the same sign of the relationship, but the estimates from the binning methods are generally greater than the MLE estimates. **NOTE** I need to re-plot these in ggplot and stitch together. Currently stitching pre-made png files, I think that's why it looks so funky.

209 The β_1 coefficient estimates for the AMD data are all positive and of relatively similar magnitudes across
210 methods. The MLE estimate has the smallest magnitude and SD (0.063 ± 0.011 SD), The ELBn method
211 is slightly larger and has a wider SD : (0.072 ± 0.018). Finally, the NAS method has the largest estimate
212 and an intermediate standard deviation (0.076 ± 0.013). The range of the gradient in the AMD data is
213 relatively small (9.5), but the absolute change in the size spectra parameter (λ) across the AMD gradient
214 vary by ~0.13 units depending on the method used (range of absolute change: 0.59 to 0.72).

215 Similar to the AMD results, the β_1 coefficient estimates for the NEON data are all negative and have
216 similar magnitudes across methods. The MLE estimate is once again the smallest (-0.006 ± 0.001 SD).
217 The NAS and ELBn methods are larger at -0.004 ± 0.002 , and -0.002 ± 0.002 , respectively. Although the
218 estimates are similar and have a small magnitude, the range of the gradient in this data is larger (~30°C),
219 meaning that the absolute change in λ estimates varies by 0.11 units depending on the methods used (range
220 of absolute change across gradient: 0.06 to 0.17).

221 Simulating shallower lambdas

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

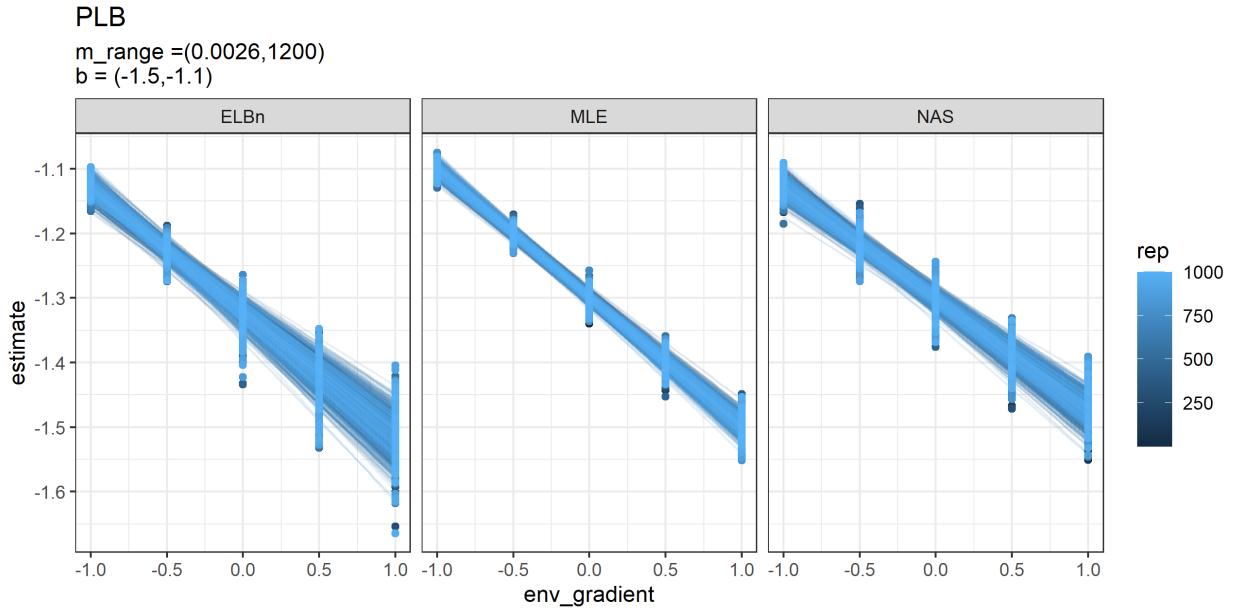


Figure 8: Relationship estimates across the hypothetical gradient for each replicate. Each panel is a different method for estimating the size spectra parameter. (match to “new” style if needed)

We found generally less variability in the relationship estimates (β_1) across a gradient of distributions described with shallower λ parameters. However, the median MLE estimate was once again closest to the hypothetical relationship and had the narrowest distribution. The median ELBn estimate was close to the hypothetical relationship but had a slightly wider distribution. Finally, 95% of the NAS estimates were greater than the hypothetical relationship. In general, the distributions from the different methodologies followed the same overall pattern, with the MLE estimates being the smallest, and the binning methods being larger (i.e., more positive).

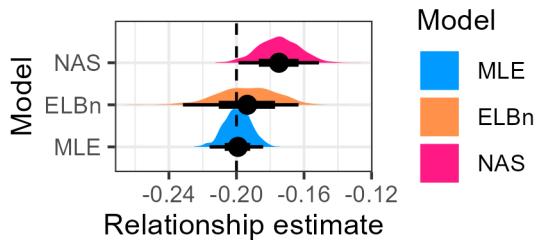


Figure 9: Distribution of relationship coefficient estimates. Vertical line is the known relationship. All methods under estimate the value, but the mean magnitude and distribution of values is greater for the ELBn and NAS methods. (not sure why this is such different proportions, need to re plot with right dimensions)

```

236 ##   Model    mean      p50      p25     p975
237 ## 1   MLE -0.199 -0.199 -0.205 -0.183
238 ## 2   ELBn -0.195 -0.195 -0.207 -0.161
239 ## 3   NAS -0.174 -0.174 -0.183 -0.151

```

NOTE Not sure if these tables are useful or not. Shows same information as figures, but might be easier to see the quantitative values instead of estimating from plot

242 Discussion

243 Measuring parameters describing the decline in abundance with increasing body size in communities
244 is being done with increasing frequency across ecology. Previous work has investigated the accuracy and
245 inherent biases associated with different estimation methods. However, how these inaccuracies and biases
246 compound across environmental gradients remains uncertain, making it difficult to detect variation in size
247 spectra parameters across environmental gradients with confidence. Here, we sampled body sizes from known
248 distributions with varying parameters (λ) and estimated the coefficient of the relationship (β_1) across a hy-
249 pothetical gradient in order to assess how the results would vary depndng on the method used. Likewise, we
250 compared how the interpretation of previously published results could change depending on the methodology
251 used to estimate the relationship.

252 Generally, the estimate from the MLE method was always closer to the known λ value, and resulted in
253 the most accurate β_1 coefficient estimate for the hypothetical relationship in simulated data, and always had
254 the smallest variation around the estimate.

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

264 The variation in estimated relationships across the empirical datasets varied by ~0.1 units depending on
265 the method used. At first glance this may not appear to be a concerning magnitude of variation. However,
266 when one considers the range of variation reported in many seminal works on size spectra relationships in
267 streams is ~0.1-0.2 (SI Table xx), it is apparent that many of these results may be dependent on the method
268 used.

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

279 Despite the drop in performance with reduced variation in λ values, the estimated relationship coefficient
280 was generally in the correct direction and of a similar magnitude. This suggests that previously reported
281 significant changes in size spectra parameters across environmental gradients and in experimental manipula-
282 tions are plausible. Given that all of the data within a study is treated identically, the the over all change in
283 size spectra parameters is likely reasonable. However, the biases and inconsistencies in relationship estimates
284 presented here suggest that it would be difficult if not impossible to compare the relative changes across
285 different published studies which use different methods. Publication of individual body size data with future
286 studies of size spectra relationships would greatly aid in our ability to generalize changes to this fundamental
287 aspect of community organization across spatiotemporal scales and in response to environmental conditions.

288 **NOTE** not sure if the idea below belongs in this paper or not, but it's something I've been thinking
289 about for a while. I'm using this as an excuse to write out my thoughts and to see what y'all think about
290 it. Maybe it's another paper to write in the future?

291 An interesting finding is that the empirical data sets used here were quite a bit shallower than is ex-
292 pected based on theory, and that what has been observed empirically in marine systems. There are two,
293 non-mutually exclusive, possible explanations for this: 1) freshwater streams are more dependent on detrital
294 resources, 2) streams are effectively 2D systems, whereas marine systems are 3D. Blanchard et al. (2009)

295 observed that benthic marine communities had shallower slopes than pelagic communities, and that this was
296 likely due to a stronger reliance on a shared detrital basal resource in the benthic community, whereas the
297 pelagic community was largely driven by energy transfer through predator-prey interactions. The impor-
298 tance of allochthonous resources in stream ecosystems has long been recognized (Vannote et al. 1980). The
299 stream data here largely comes from low order, headwater streams which may be reasonably considered to
300 be driven by allochthonous resources. It is possible that the communities sampled here are structured based
301 on a shared basal resource, and that the importance of predator-prey interactions is reduced. An alternate
302 hypothesis is the relative dimensionality of search space in benthic and pelagic systems. Consumption rates
303 and interaction strengths have been shown to vary markedly based on search dimensionality. Consumers
304 generally encounter resources more frequently in 3D environments (i.e., aerial, pelagic) than in 2D environ-
305 ments (i.e., benthic, terrestrial), and lead to stronger consumption rates and interaction strengths. Likewise,
306 it has been hypothesized that the scaling exponent for mass-specific search rates may vary based on search
307 space dimensionality instead of being set at approximately 0.75 (Pawar, Dell, Savage, 2012).

308 Concluding Remarks

309 Both of the binning methods can perform as well as MLE under certain circumstances (large sample
310 size, shallow λ). However, it would be unlikely for a researcher to know if this was the case beforehand.
311 Furthermore, in order to confirm this result, MLE methods would likely need to be employed. With the
312 publication of the *sizespectra* package (2017), producing MLE estimates of size spectra parameters is a
313 relatively easy task. Therefore, we recommend using this in all future studies of size spectra relationships.

314 We reiterate the recommendations of White et al. (2007)(2007), Sprules and Barth (2016)(2016) and
315 Edwards et al. (2017)(2017) to estimate size spectra relationships using MLE methods due to their superior
316 performance in nearly every context. Furthermore, we strongly encourage authors to publish individual size
317 data whenever possible. This will allow for the consistent re-analysis of existing data sets as methodologies
318 develop and improve. This will aid in the ability for size spectra work to be synthesized between research
319 groups and across scales.

320 **References**

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

³³⁹ **Supplementary material**

³⁴⁰ **Range of body sizes, M**

³⁴¹ **NOTE** update all plots to be consistent with main text

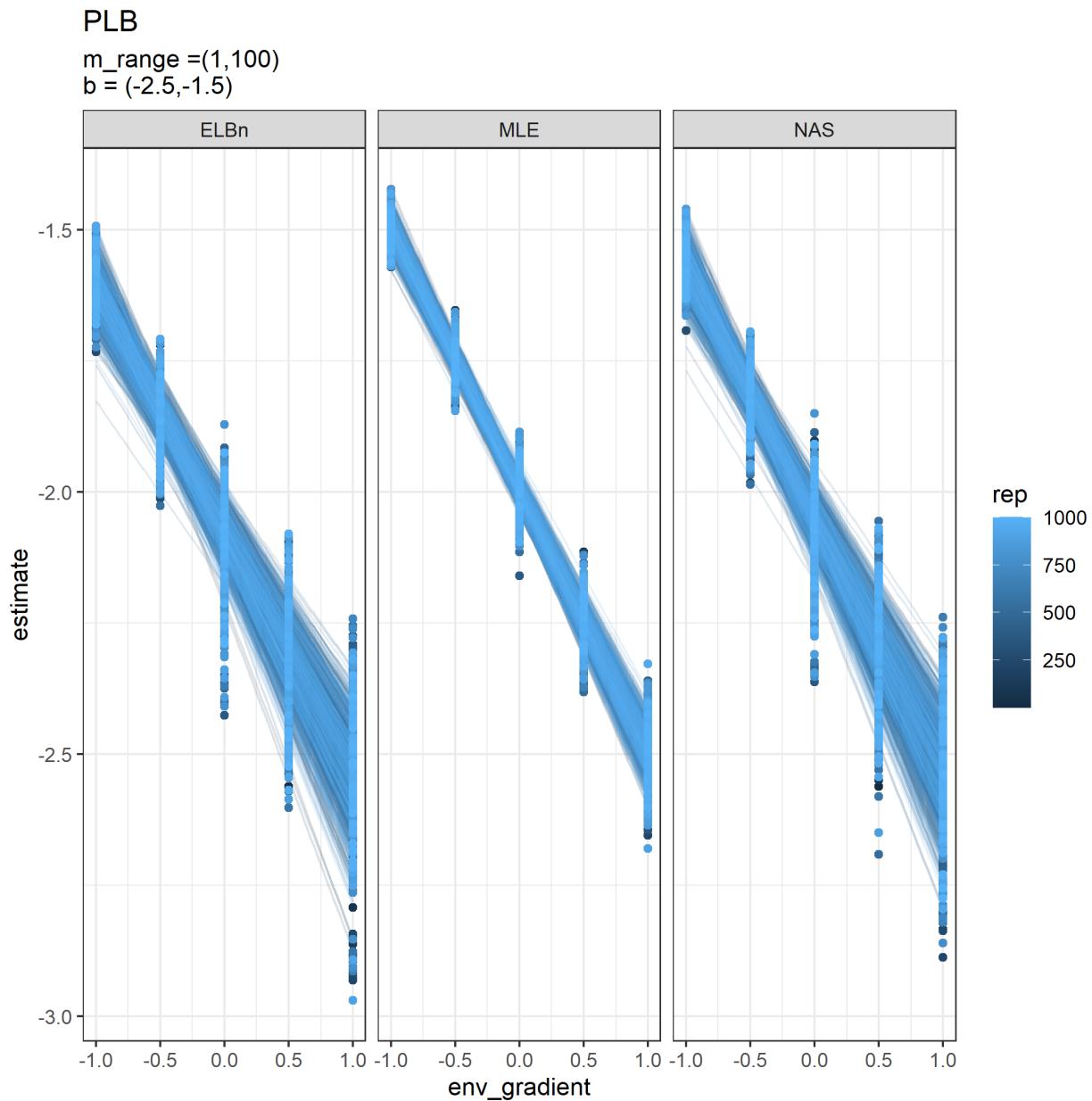


Figure S1: Individual regressions for five sites across a hypothetical gradient with a known relationship of 0.5. Range of body sizes is reduced and is from 1, to 100.

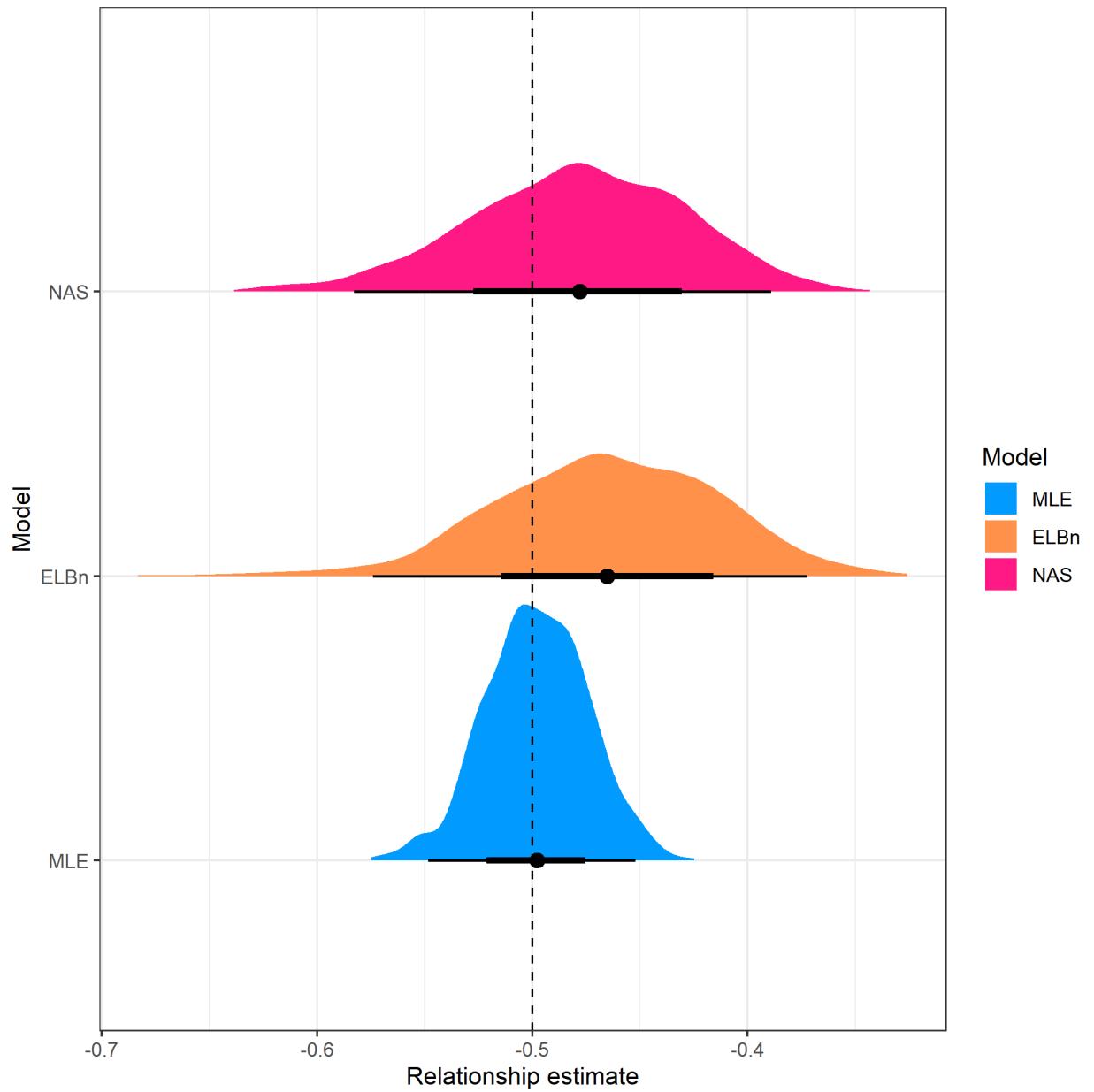


Figure S2: Distribution of estimated relationship (β_1) coefficient's for five sites across a hypothetical gradient with known value of 0.5. Range of body sizes is reduced and is from 1, to 100.

³⁴² Large environmental gradient

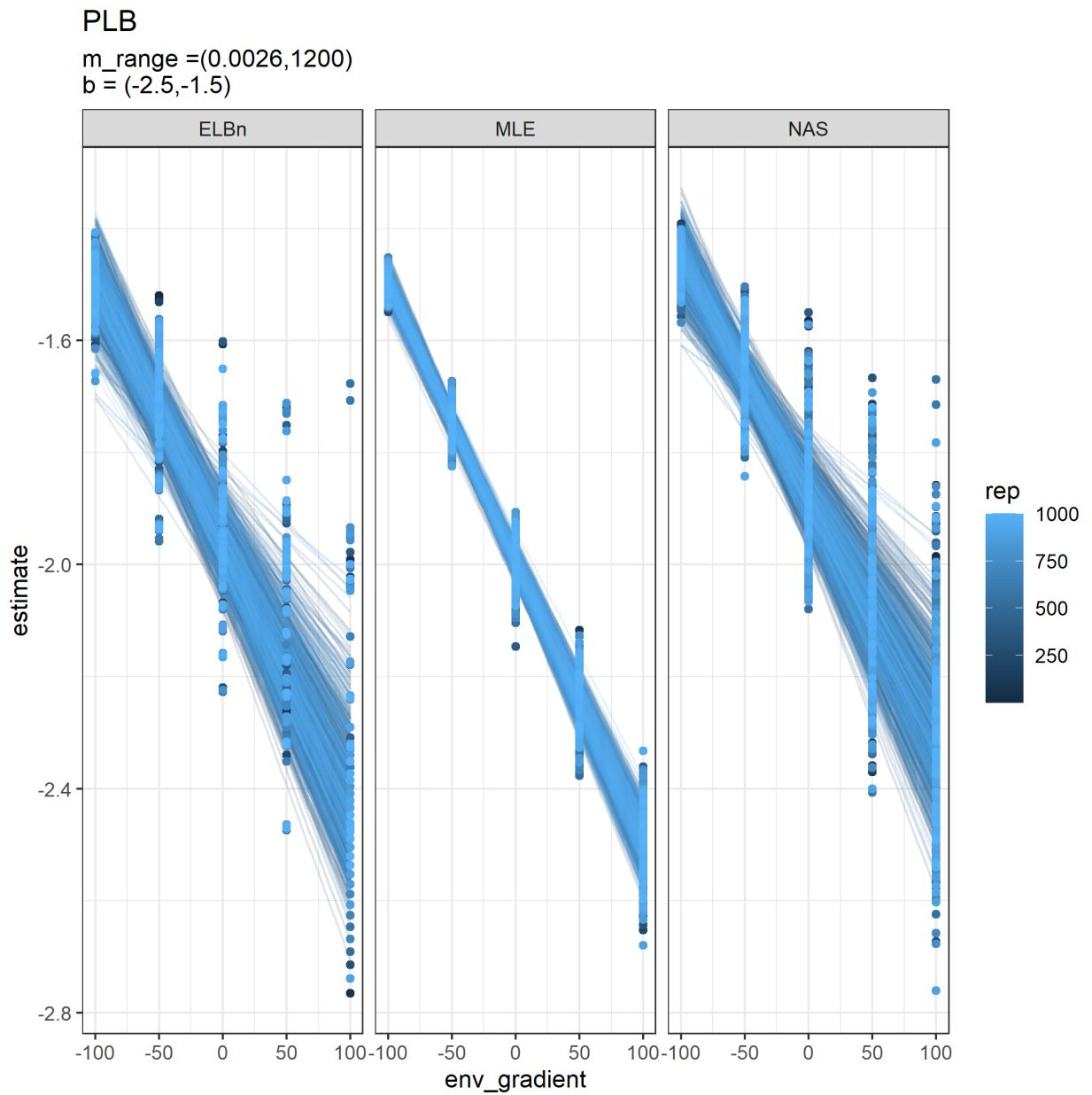


Figure S3: Individual regressions for five sites across a hypothetical gradient with a known relationship of 0.5. Range of environmental values (x -axis) increased to be -1000, to 1000.

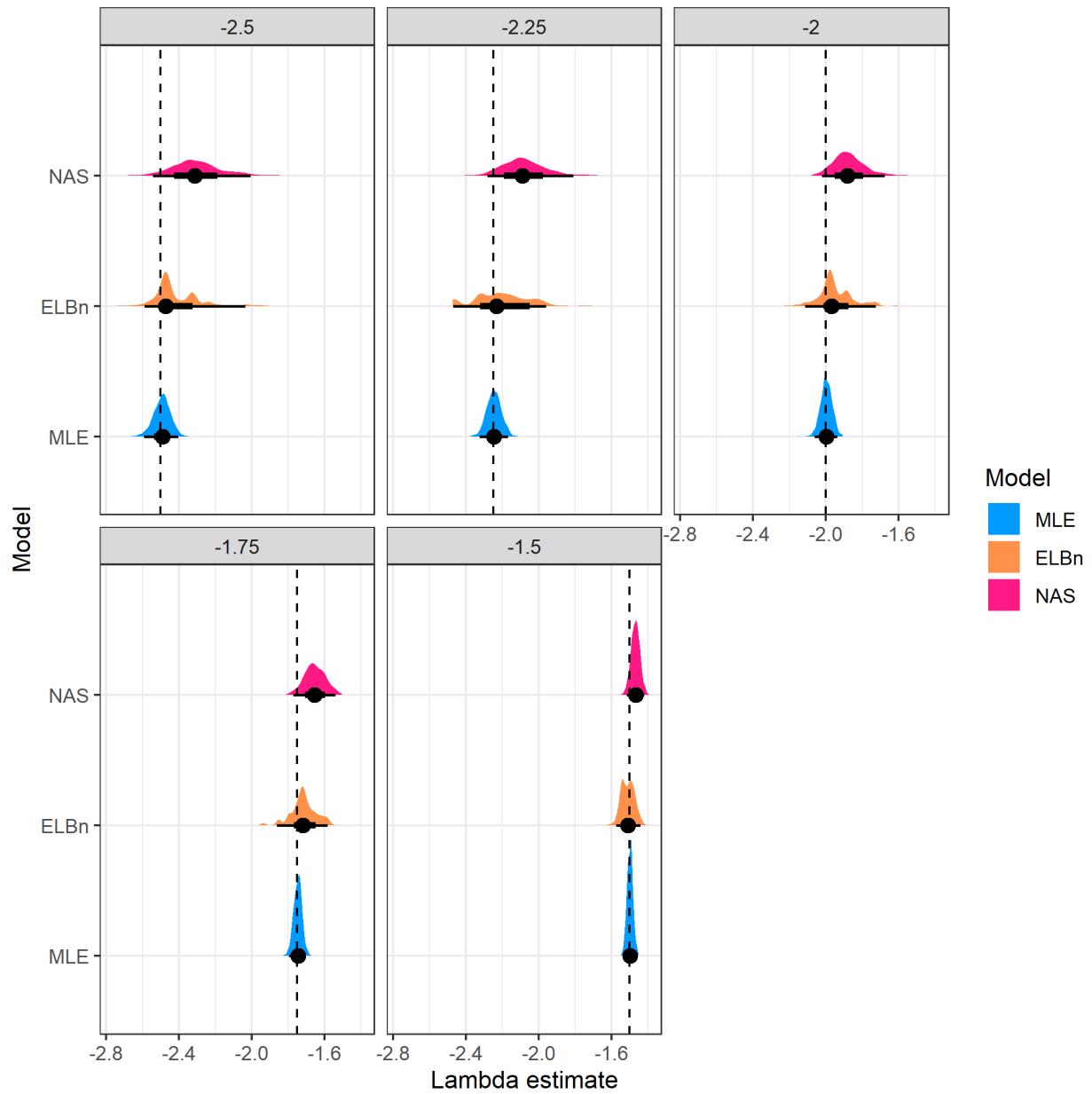


Figure S4: Distribution of estimated λ coefficient for five sites across a hypothetical gradient with known values. Range of environmental values (x -axis) increased to be -1000, to 1000.

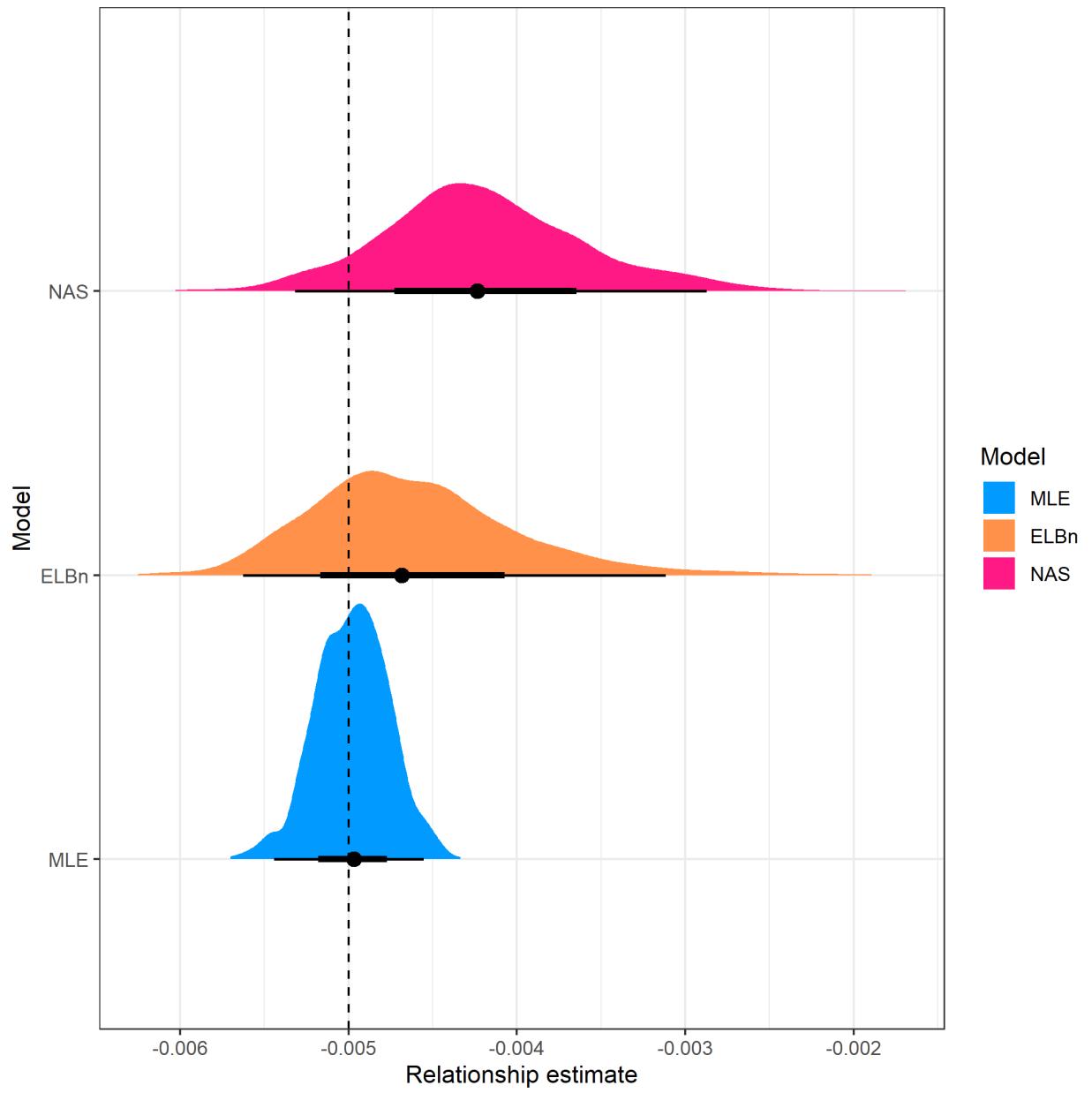


Figure S5: Distribution of estimated relationship (β_1) coefficient's for five sites across a hypothetical gradient with known value of 0.5. Range of environmental values (x -axis) increased to be -1000, to 1000.

³⁴³ Varying number of sites

³⁴⁴ 10 sites

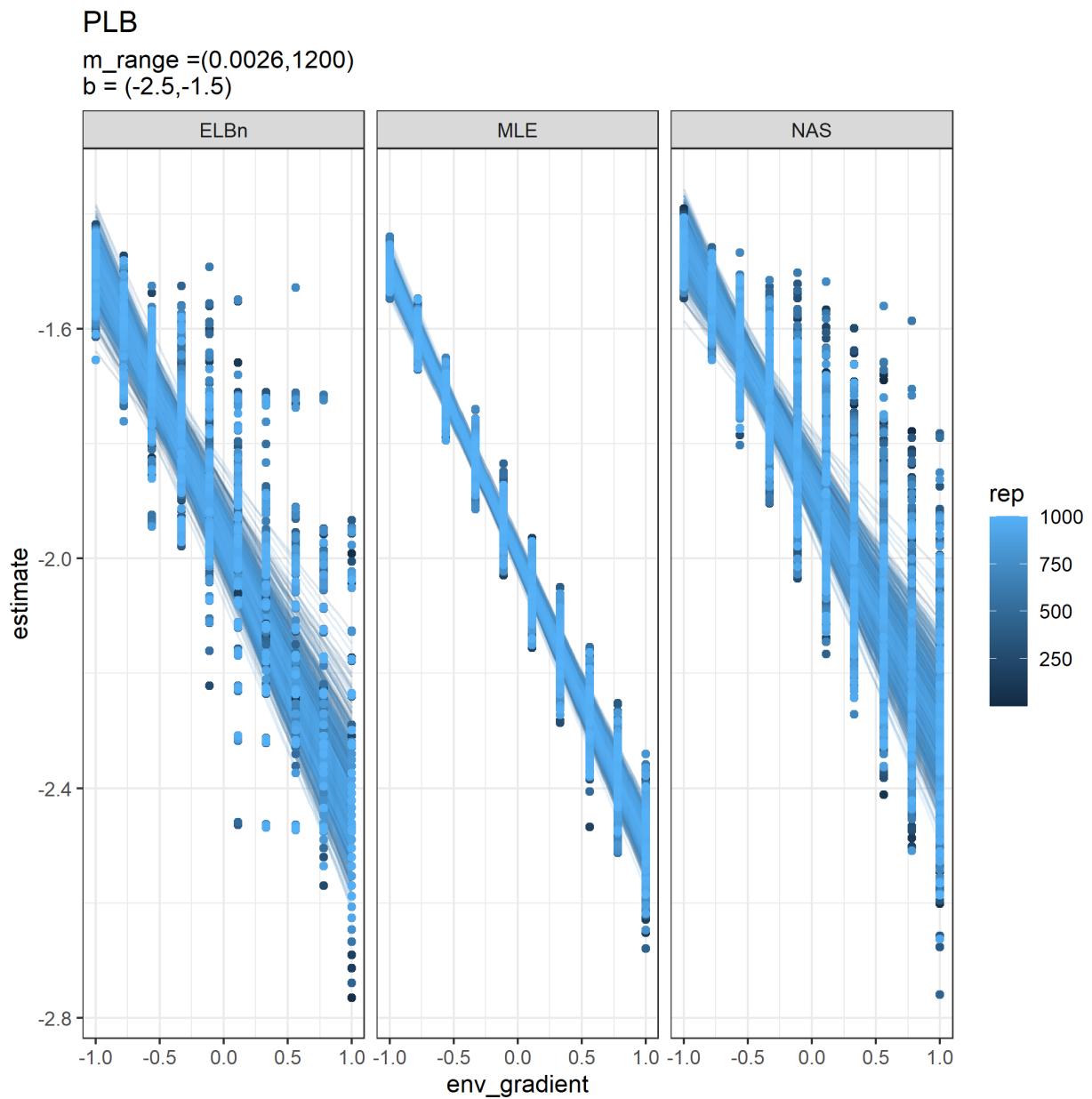


Figure S6: Individual regressions for ten sites across a hypothetical gradient with a known relationship of 0.5. All other parameters are the same as in the main analysis

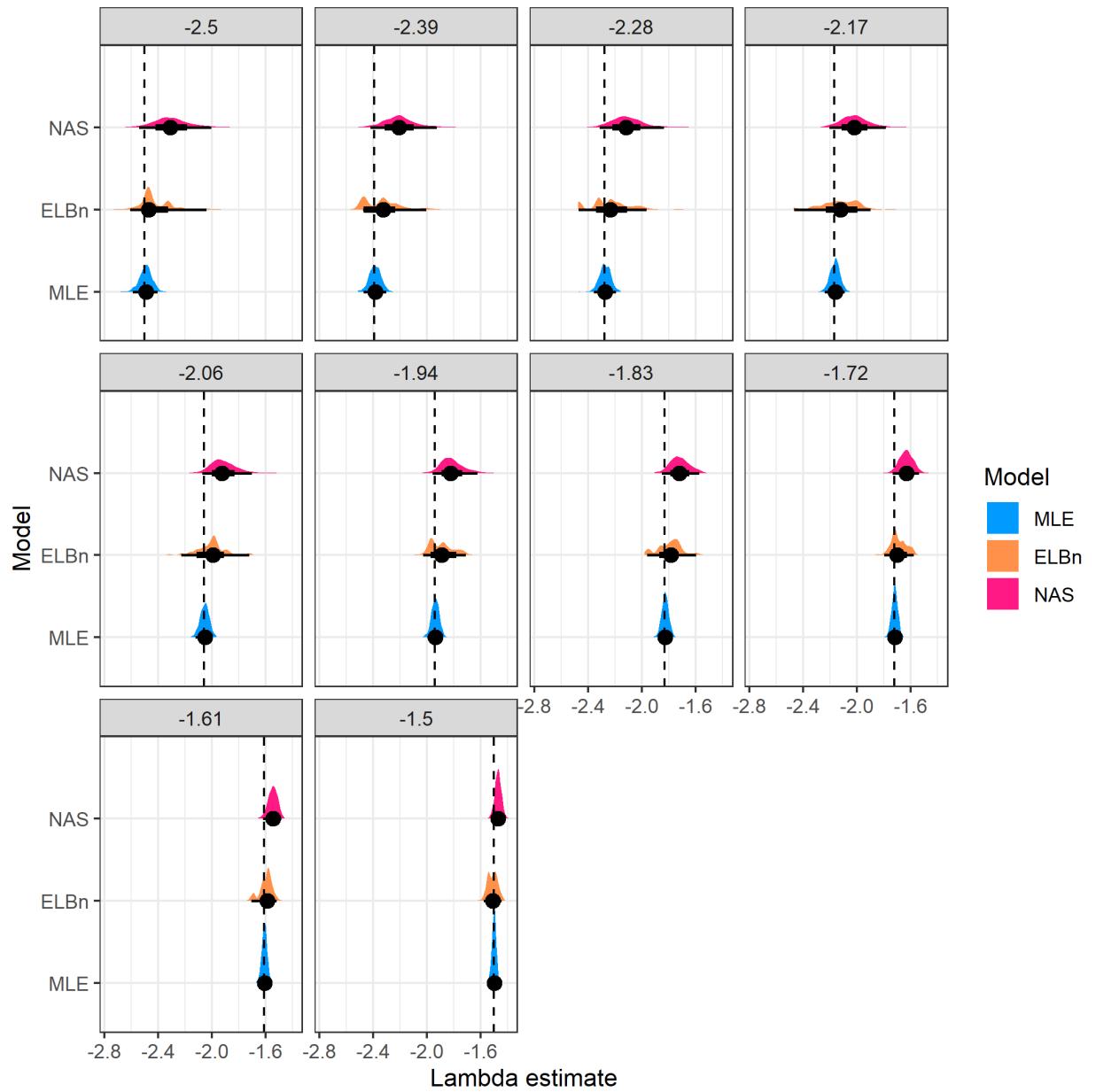


Figure S7: Distribution of estimated λ coefficient for ten sites across a hypothetical gradient with known values. All other parameters are the same as in the main analysis

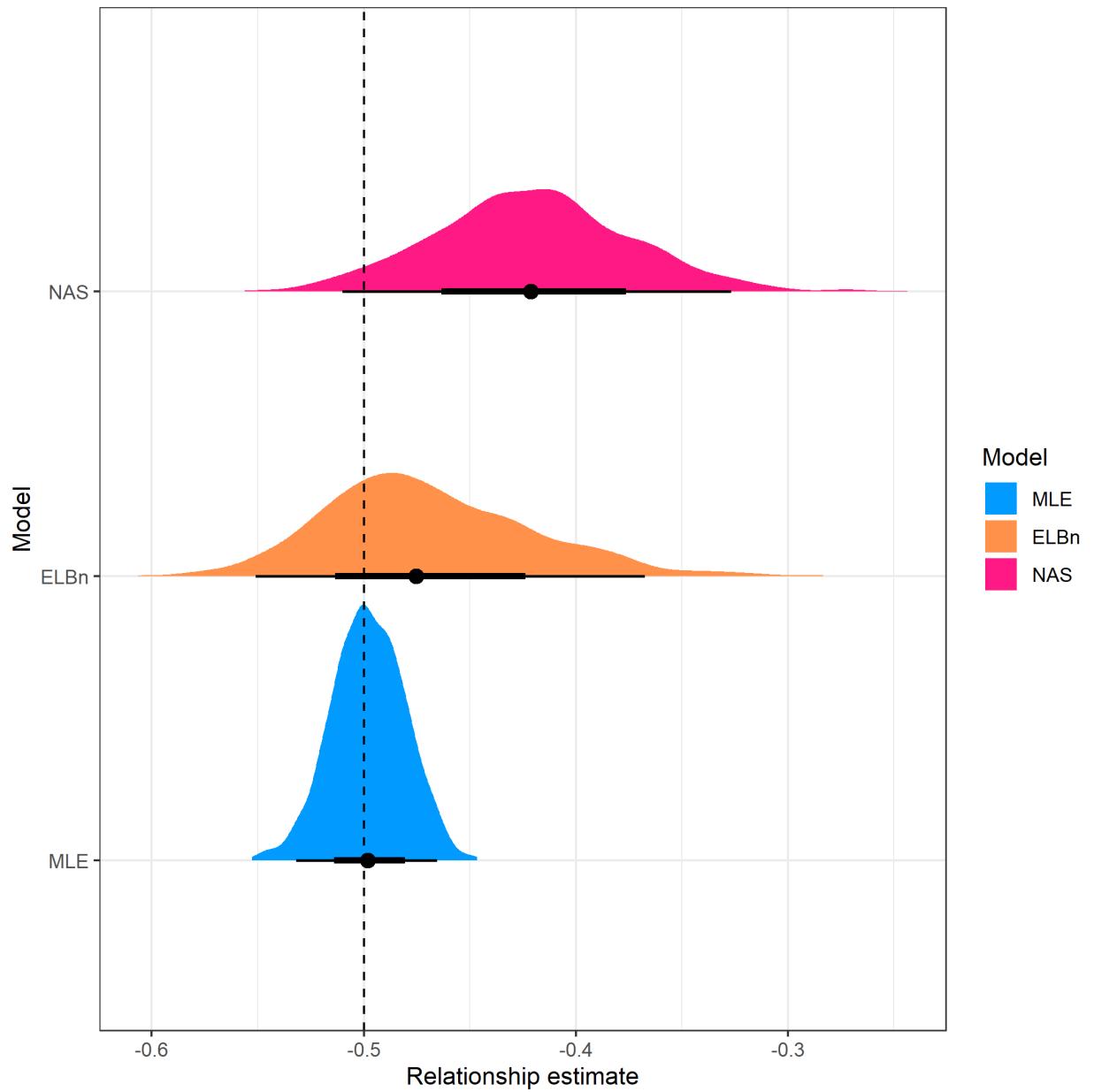


Figure S8: Distribution of estimated relationship (β_1) coefficient's for ten sites across a hypothetical gradient with known value of 0.5. All other parameters are the same as in the main analysis

PLB

m_range = (0.0026, 1200)

b = (-2.5, -1.5)

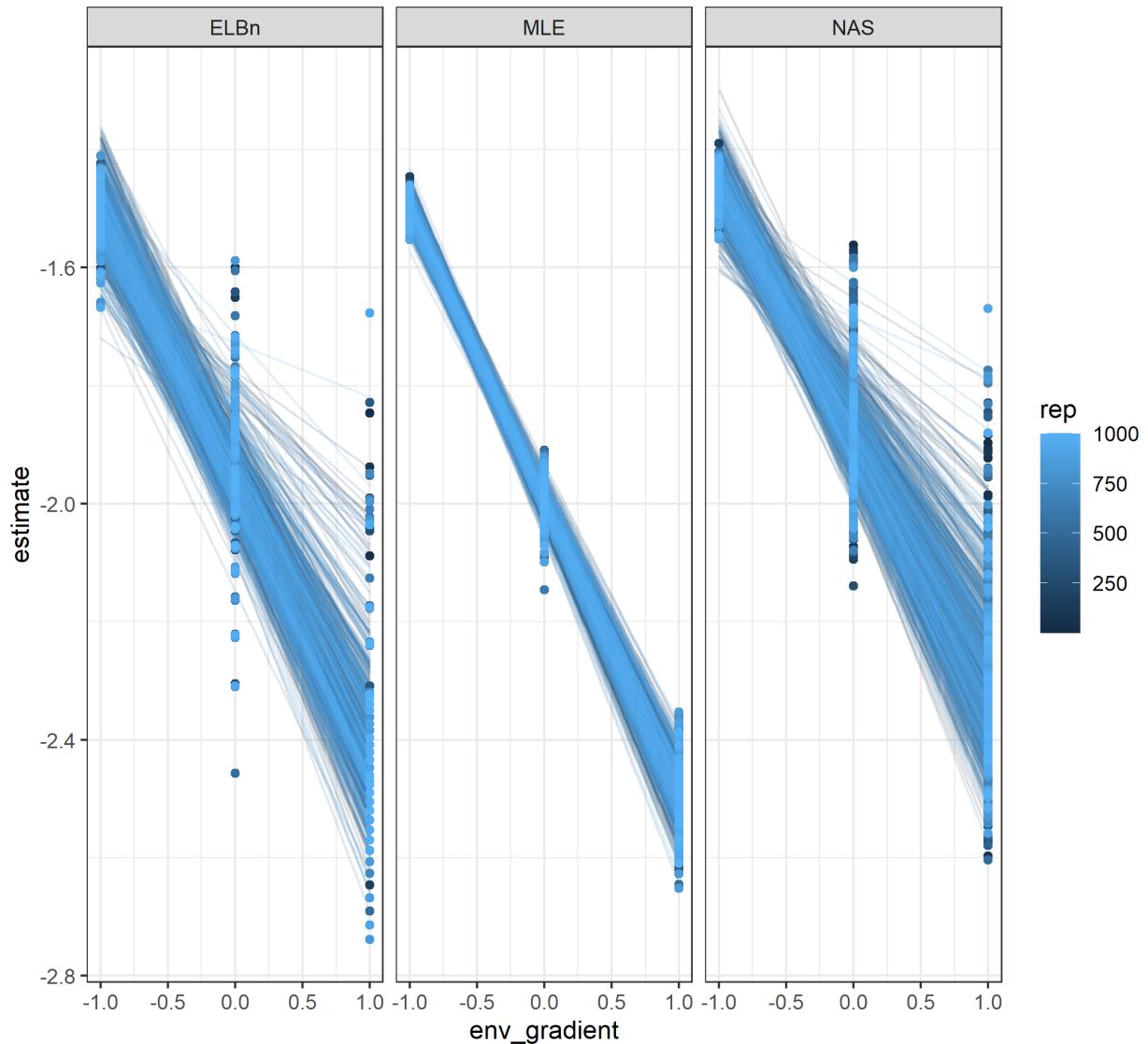


Figure S9: Individual regressions for three sites across a hypothetical gradient with a known relationship of 0.5. All other parameters are the same as in the main analysis.

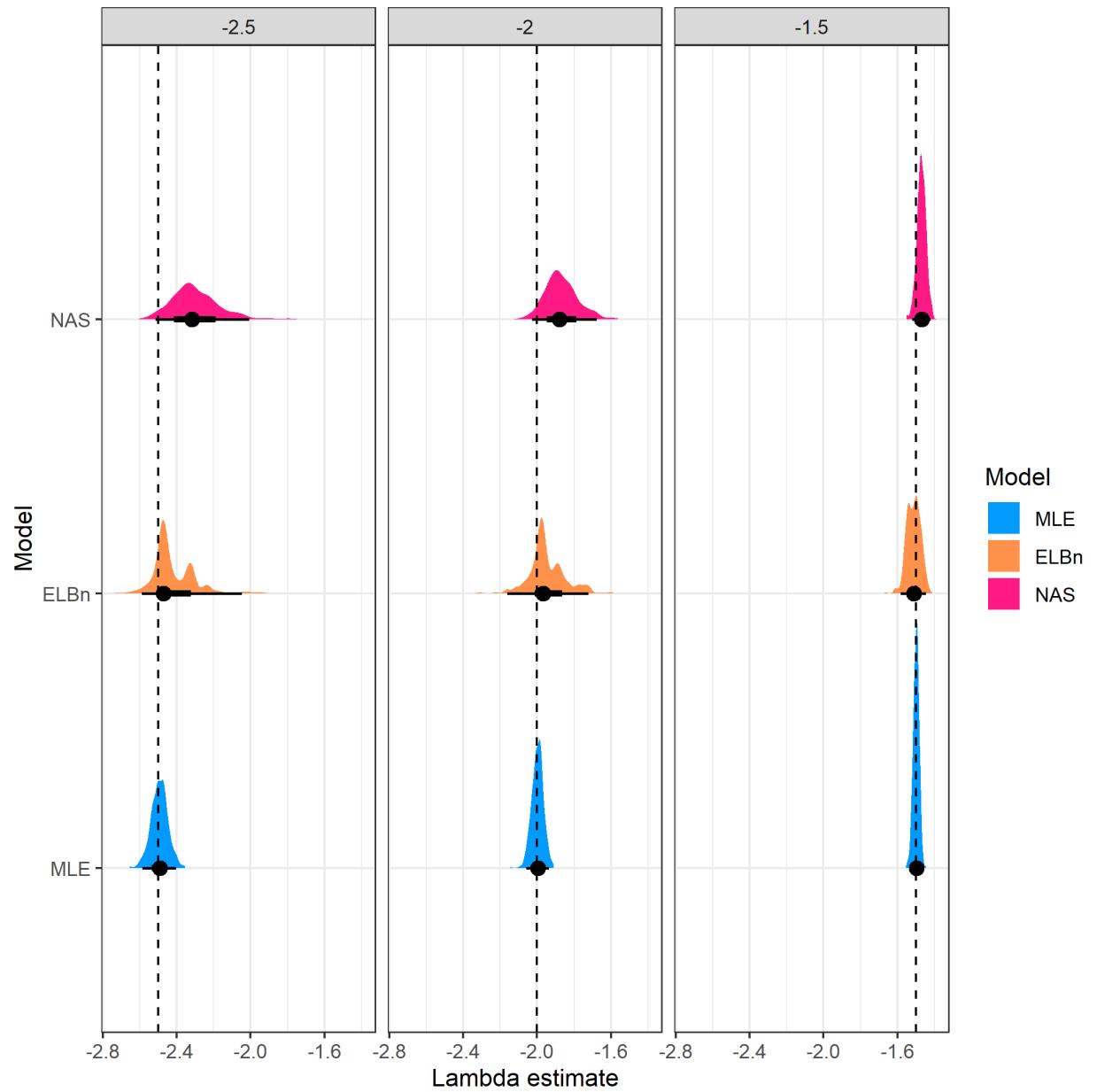


Figure S10: Distribution of estimated λ coefficient for three sites across a hypothetical gradient with known values. All other parameters are the same as in the main analysis.

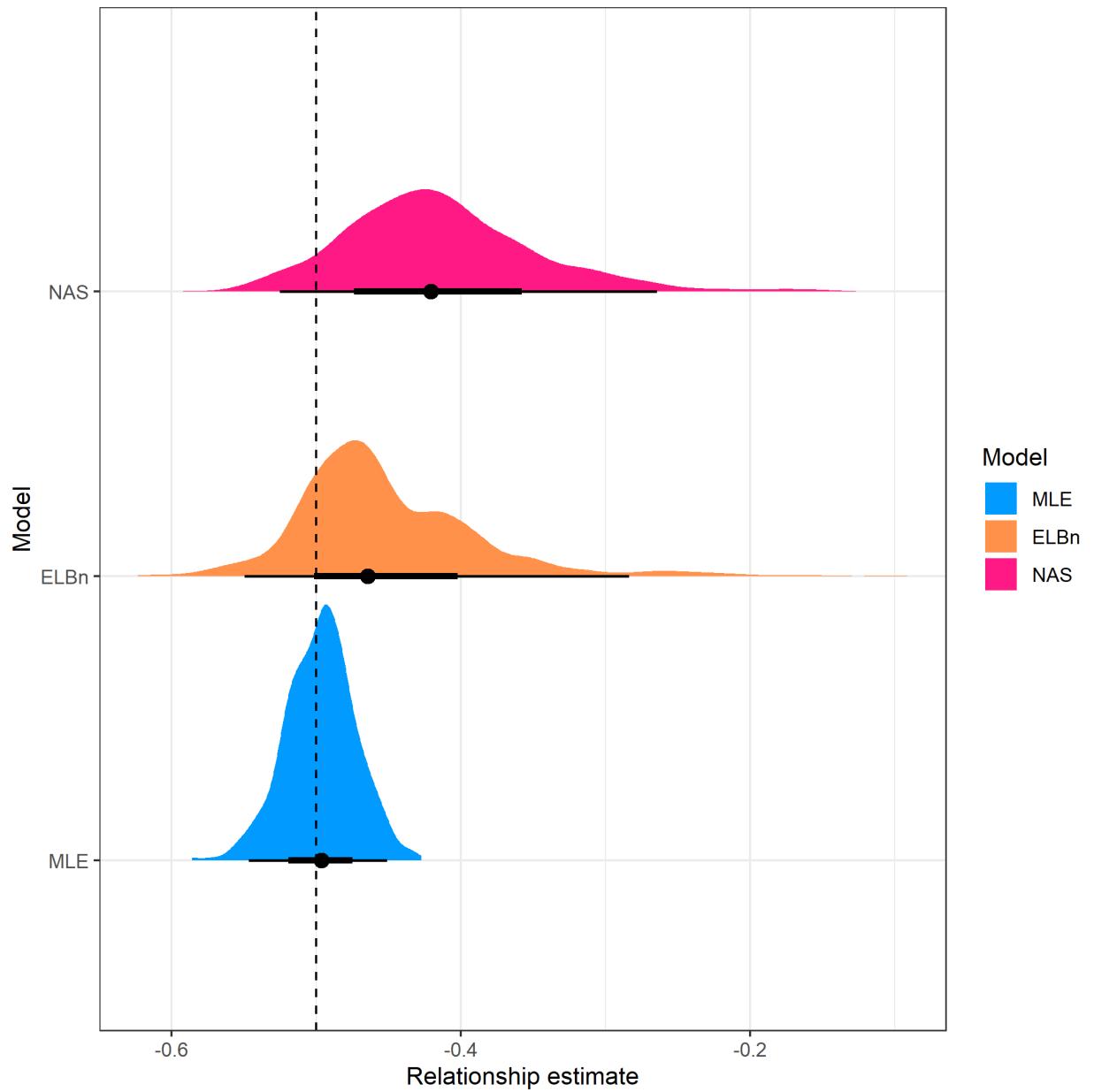


Figure S11: Distribution of estimated relationship (β_1) coefficient's for three sites across a hypothetical gradient with known value of 0.5. All other parameters are the same as in the main analysis

346 Sample size

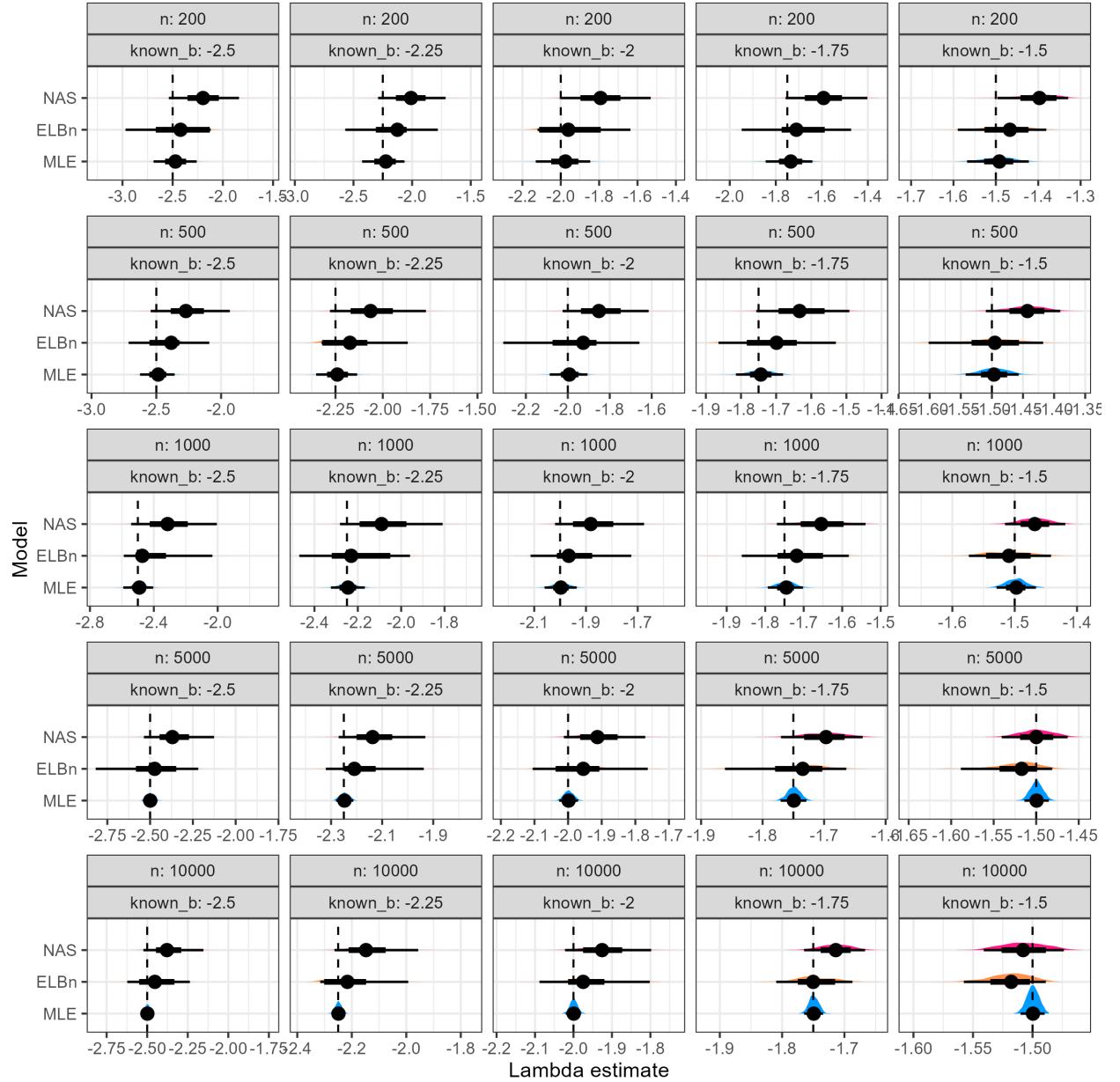


Figure S12: Distribution of size spectra parameter estimates. Vertical line is the known parameter (dashed line) which describes the bounded power law distribution from which the body size estimates were sampled. As n increases (top to bottom) and λ increases (left to right), the accuracy of the estimate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

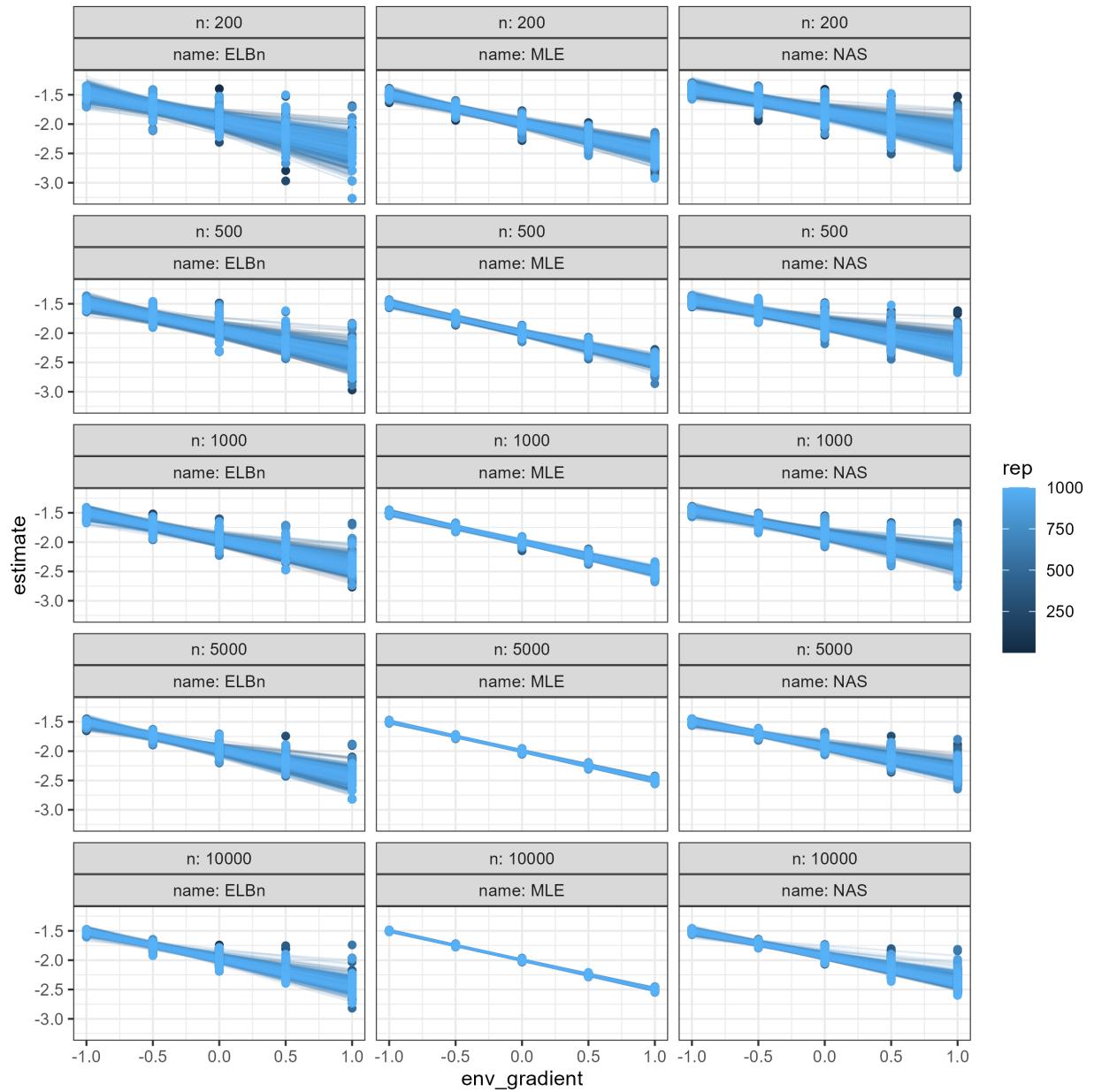


Figure S13: Individual regressions when varying sample size

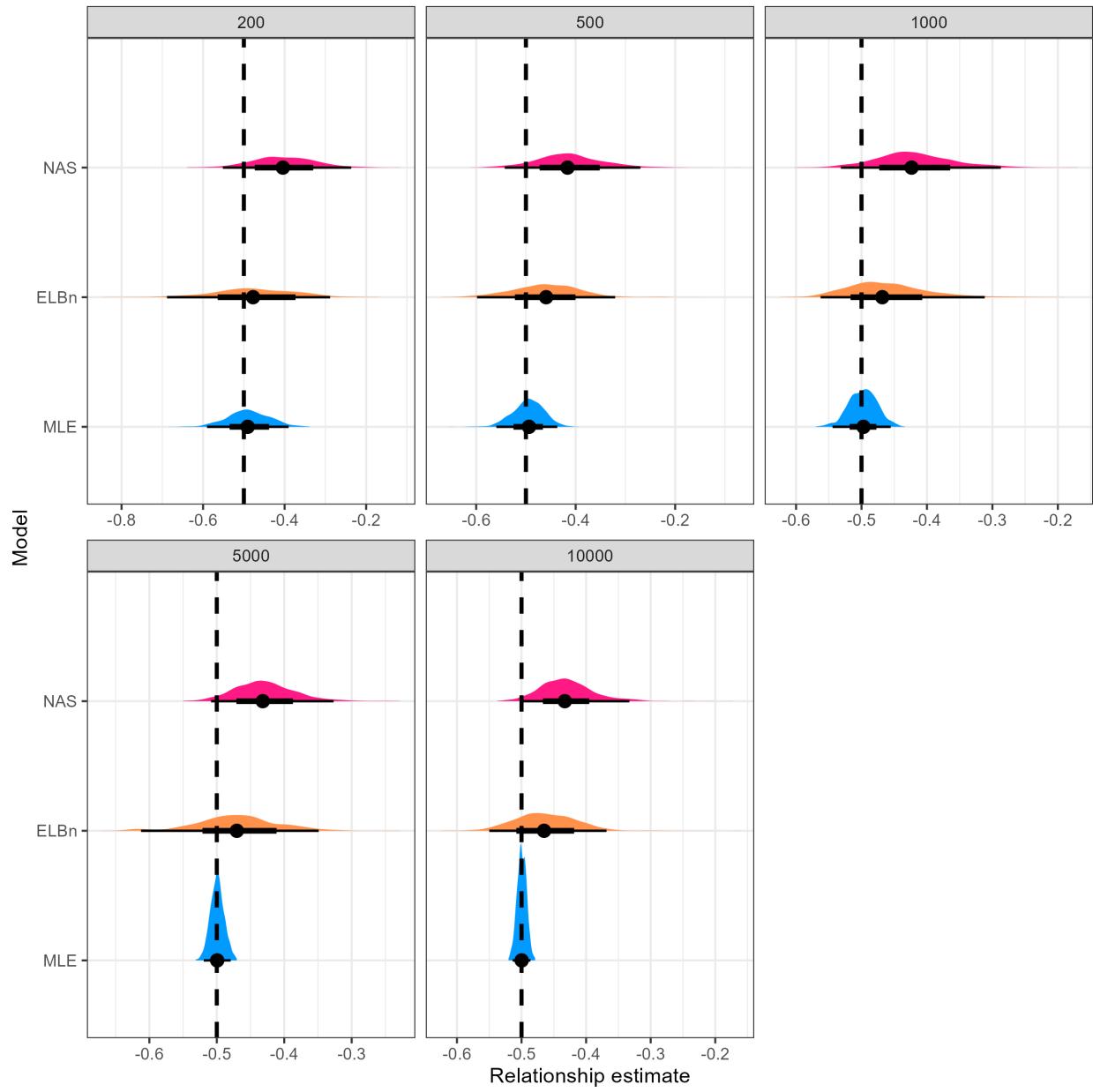


Figure S14: Distribution of the relationship coefficients with varying sample size

347 **λ and relationship estimates**

348 **NOTE** I'm considering dropping the λ estimate plots from the manuscript. I think it's too much info,
349 but if it adds to the story I'm happy to leave it in. **END**

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

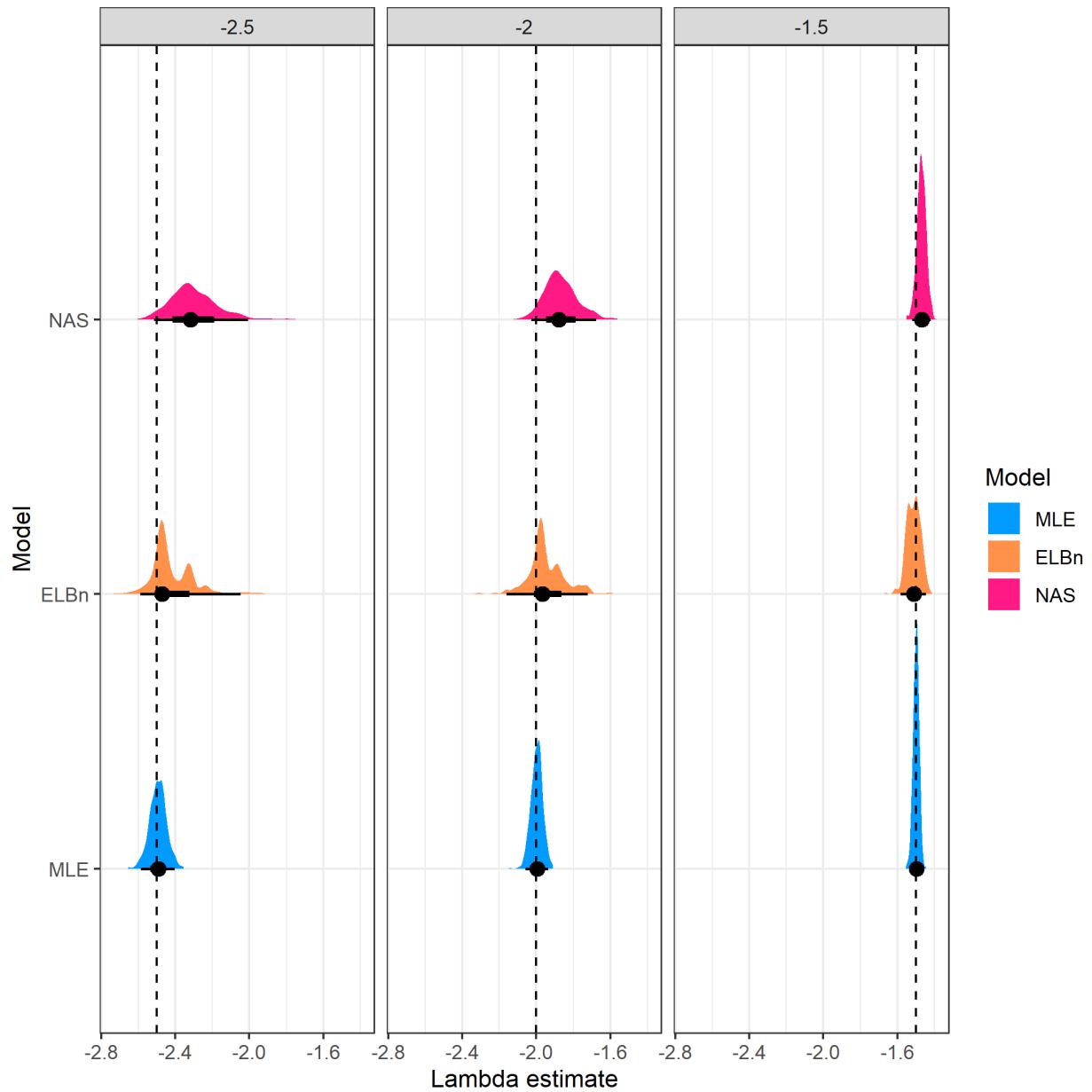


Figure S15: Distribution of estimated λ coefficient across a hypothetical gradient with known values (dashed line). The number of sites is reduced to 3, compared with 5 sites presented in the main analysis.

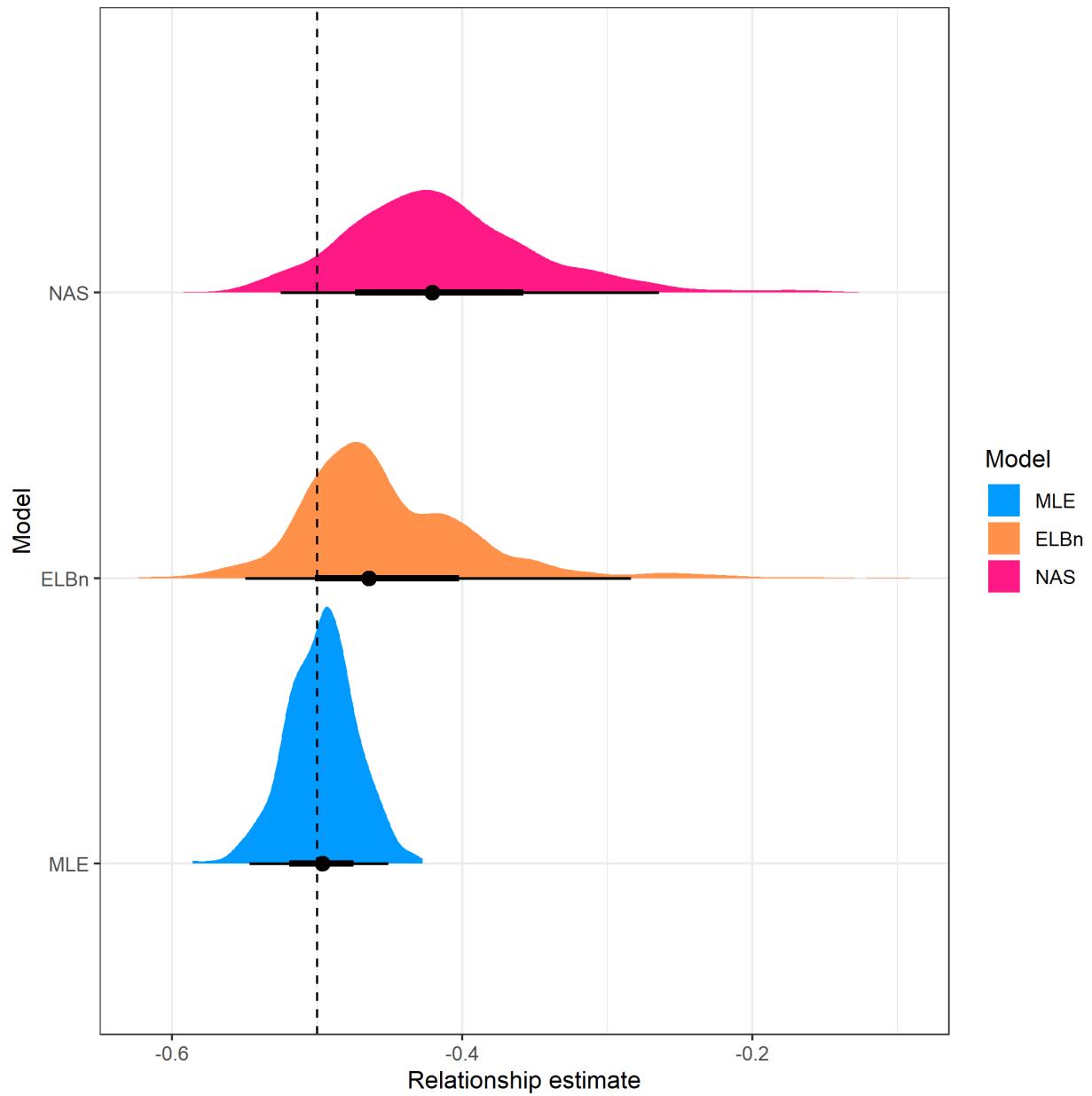


Figure S16: Relationship estimate

359 Number of sites = 10

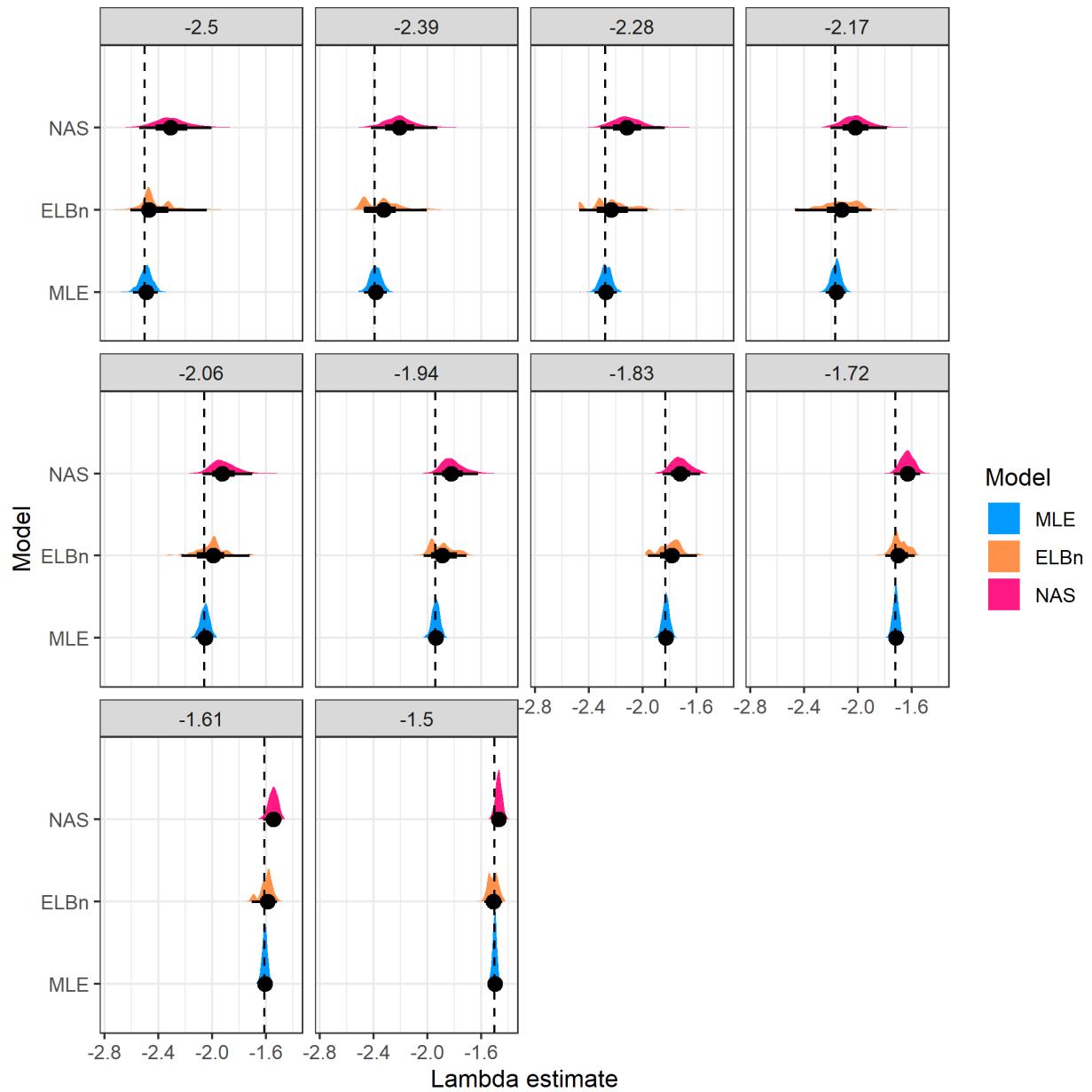


Figure S17: Distribution of estimated λ coefficient across a hypothetical gradient with known values(dashed line). The number of sites is increased to 10, compared with 5 sites presented in the main analysis.

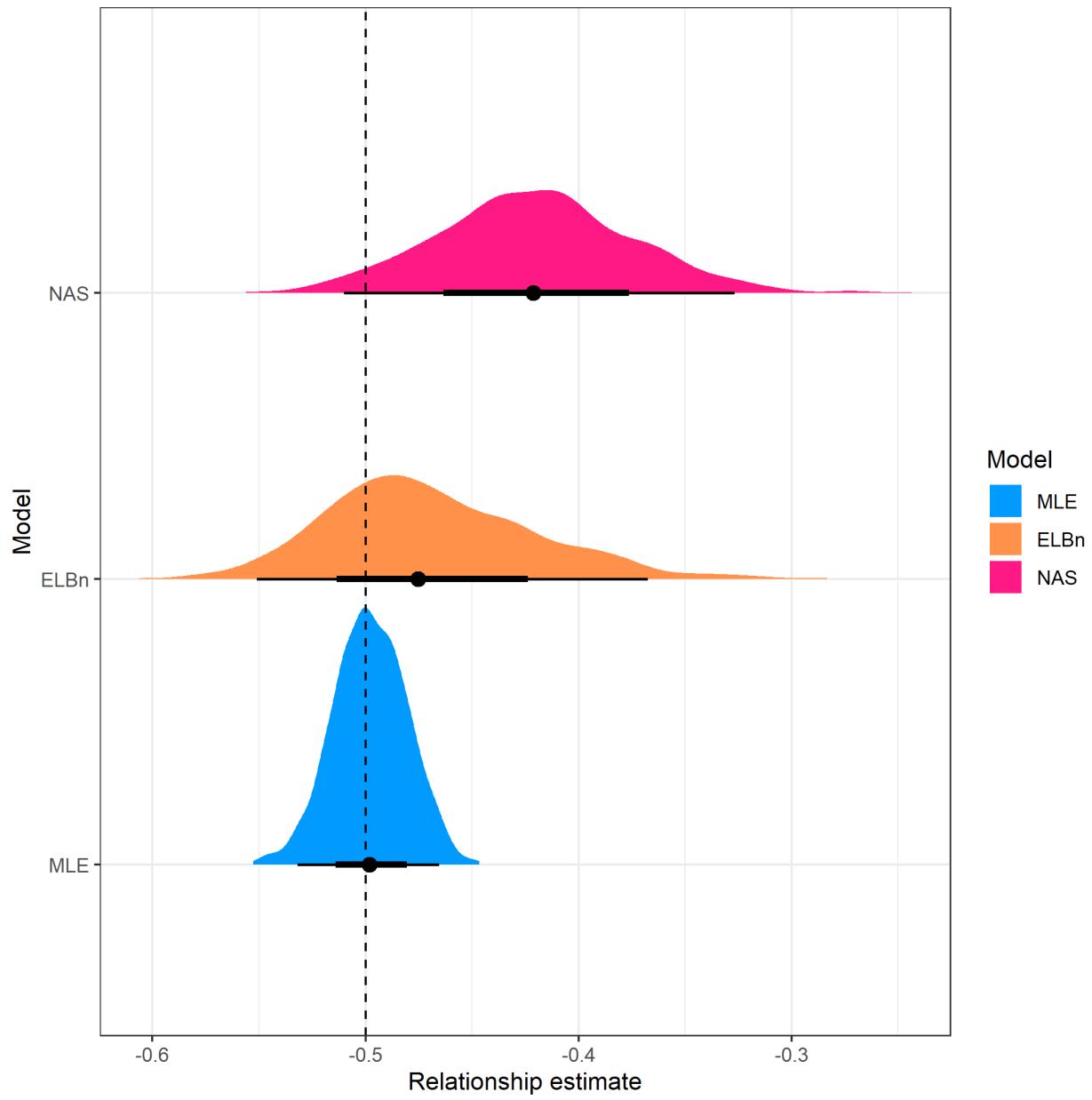


Figure S18: Relationship estimate

360 Environmental Gradient

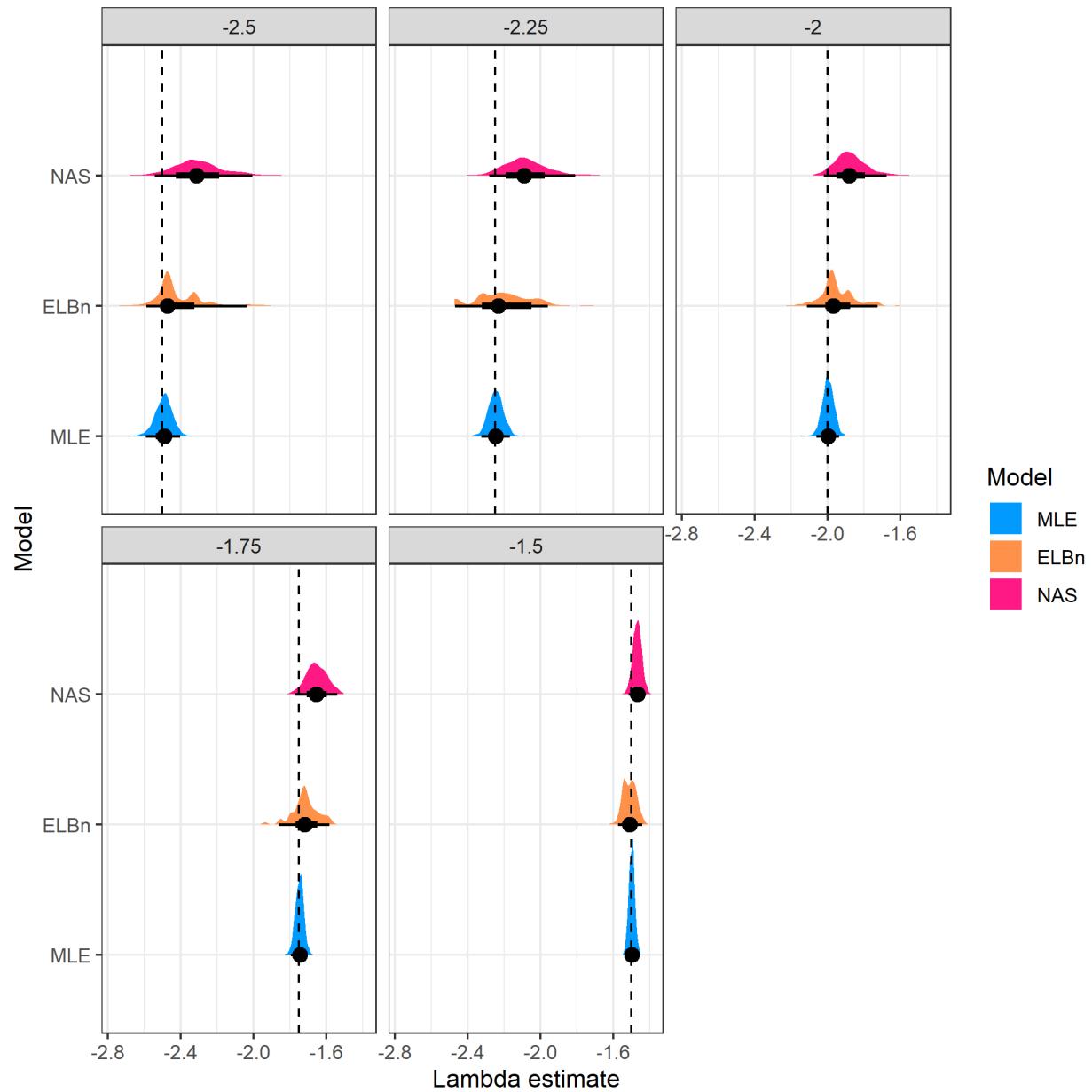


Figure S19: Distribution of estimated λ coefficient for five sites across a hypothetical gradient with known values (dashed line). The range of the environmental gradient was increased (-100 to 100).

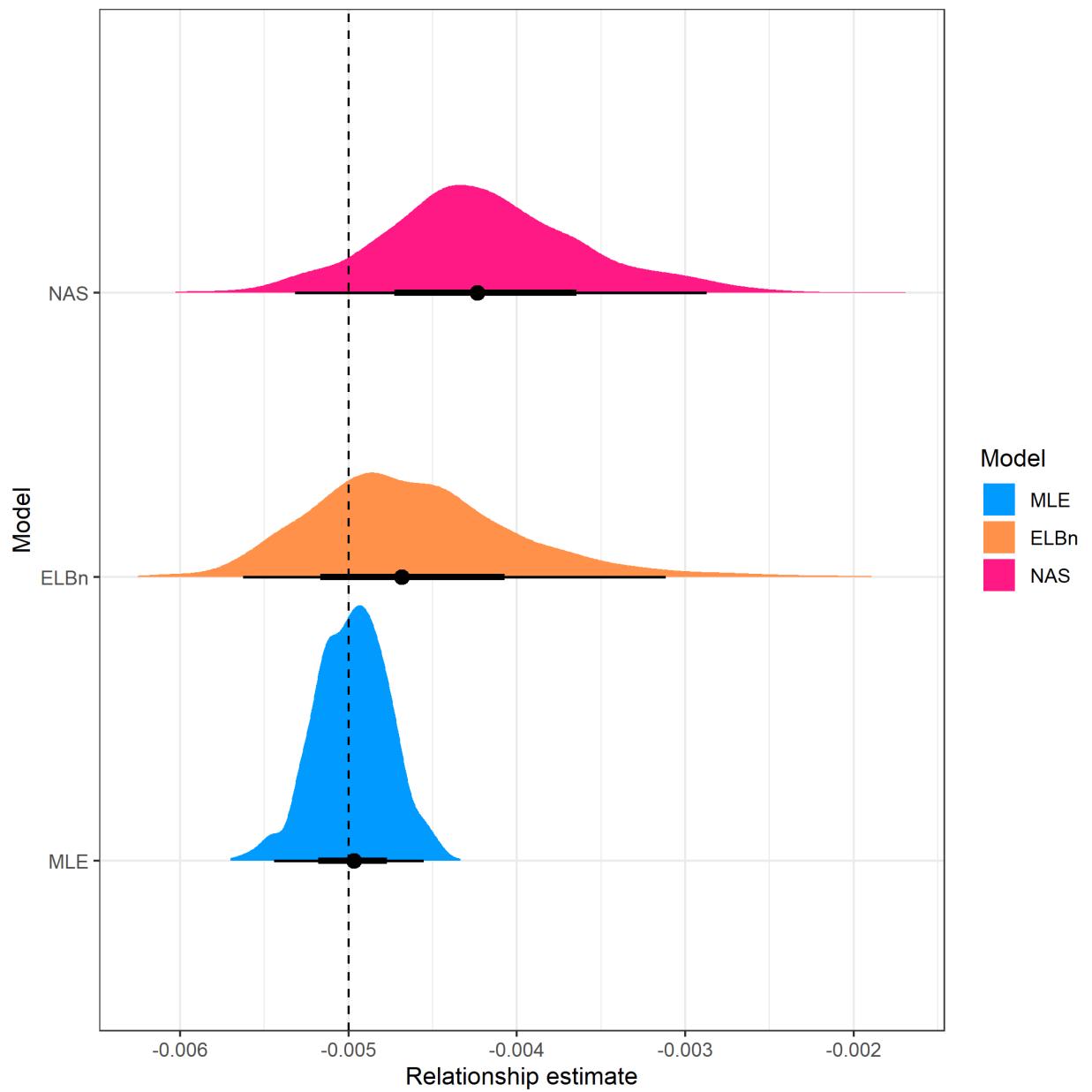


Figure S20: Relationship estimate

361 Range of body sizes = 1 to 100

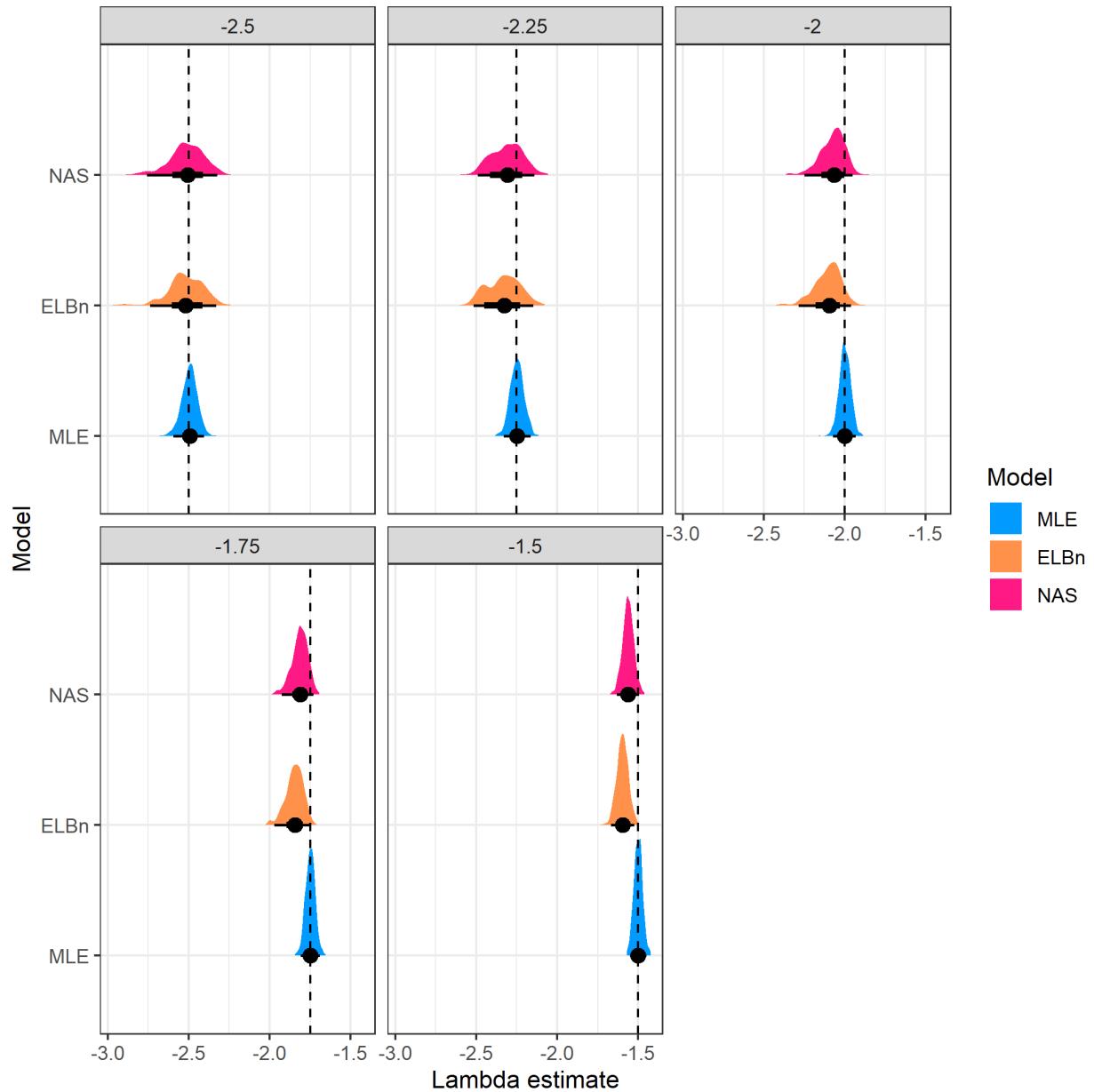


Figure S21: Distribution of estimated λ coefficient for five sites across a hypothetical gradient with known values(dashed line). Range of body sizes is smaller than main analysis and ranges from 1, to 100.

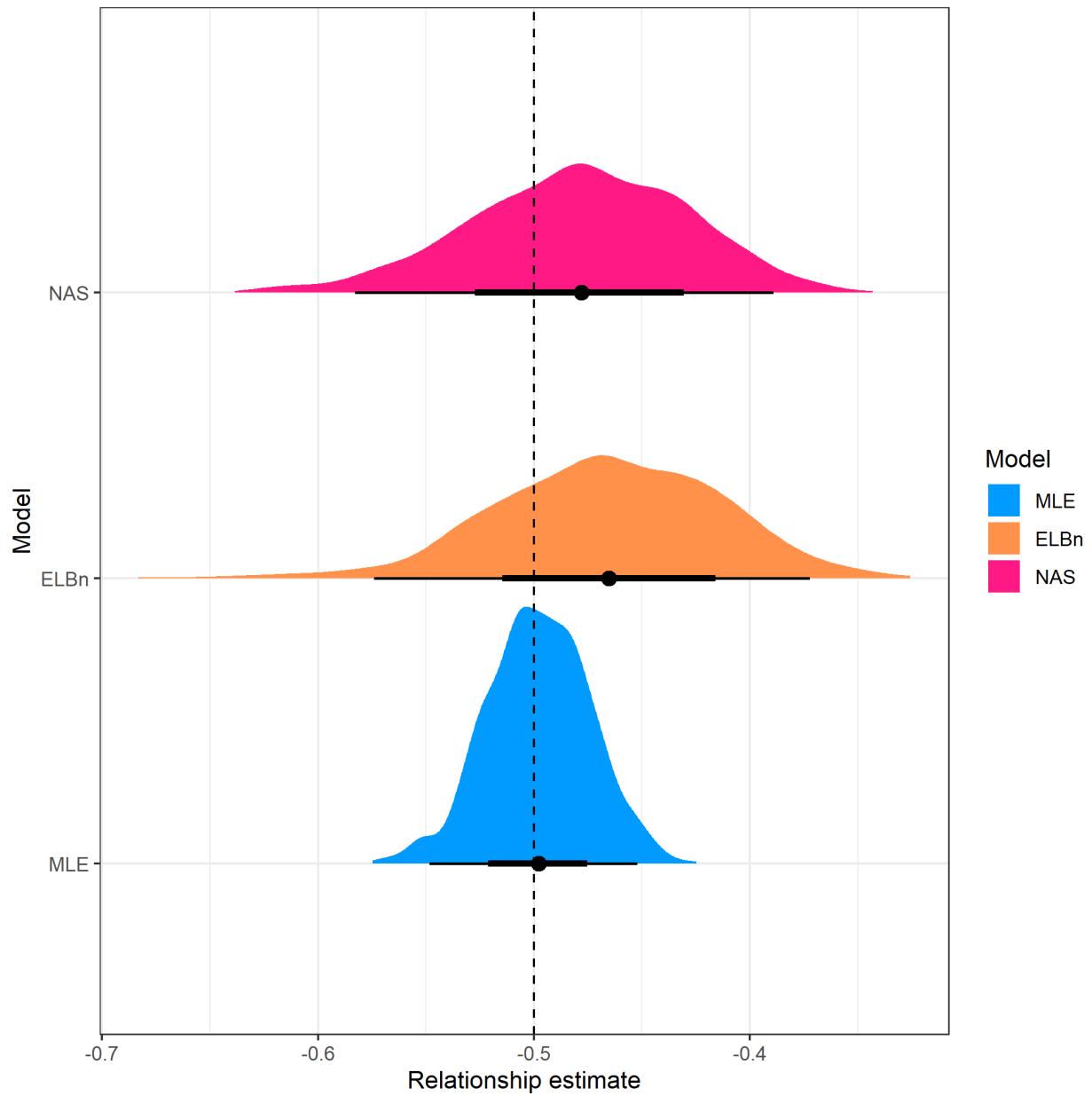


Figure S22: Relationship estimate

362 Comparison with other published estimates

363 SI Table. This table shows published estimates of the variation in size spectra slopes (or exponents) in em-
 364 pirical studies. It is unclear how to directly compare estimates of the slope with different methods. However,
 365 the published estimates here range from ~0.1 to 0.2 across the gradients studied. For comparison, the 2.5–95%
 366 quantiles around the relationship estimate for the MLE method were ~0.1, whereas for the ELBn and NAS
 367 method they were ~0.25 and ~0.2, respectively. b_{diff} is the change in estimate ($b_{\text{low}} - b_{\text{min}}$). System
 368 refers to stream communities or mesocosm experiments. Method: MLE = maximum likelihood estimate,
 369 ELB = equal logarithmic binning, the number before indicates the number of bins used. The normalization
 370 process shifts the estimates by an absolute value of 1.0. Hence, direct comparison the relative change in

371 normalized and non-normalized studies should not introduce any bias. The O'Gorman et al. 2017 study used
 372 average species size and abundance (Local Size Density Relationship, *sensu* White et al. 2007) as opposed to
 373 individual size distribution. These methods are related, but it is unclear how to directly compare estimates
 374 from each method.

```
read.csv("lit_comparison.csv")
```

```

375 ##                                     Author b_diff error error_type b_low
376 ## 1          Pomeranz et al. 2021   0.12    NA      -1.37
377 ## 2 Yvon Durocher et al. 2011 (Community) 0.13    NA      -0.79
378 ## 3 Dossena et al. 2012 (October)  0.14 0.070      se    NA
379 ## 4          O'Gorman et al. 2017   0.15    NA      -0.85
380 ## 5          Martinez et al. 2016   0.15    NA      -1.11
381 ## 6 Yvon Durocher et al. 2011 (Phytoplankton) 0.19    NA      -0.46
382 ## 7          McGarvey and Kirk 2018  0.19    NA      -1.81
383 ## 8 Dossena et al. 2012 (April)    0.21 0.064      se    NA
384 ## b_high     System     Driver
385 ## 1 -1.25     Streams Temperature
386 ## 2 -0.92 FW mesocosms Temperature
387 ## 3 NA FW mesocosms Temperature
388 ## 4 -0.70     Streams Temperature
389 ## 5 -0.96     Streams Land Use
390 ## 6 -0.65 FW mesocosms Temperature
391 ## 7 -1.62     Streams Seasonality
392 ## 8 NA FW mesocosms Temperature
393 ##
394 ## 1                               Method
395 ## 2                               MLE
396 ## 3
397 ## 4 ln (mean species abundance) ~ ln(mean species mass)
398 ## 5
399 ## 6
400 ## 7
401 ## 8
```