Manuscript: Detecting differences in Size Spectra

Justin Pomeranz1,✉, James R. Junker2,3, Vojsava Gjoni4, and Jeff S. Wesner4

11 January, 2023

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](mailto:jfpomeranz@gmail.com)>](mailto: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.

# Abstract

1. The distribution of body size in communities is remarkably consistent across habitats and taxa and can be represented by size spectra relationships.
2. Many methods have been proposed for constructing size spectra relationships, but recent work has shown that varied methods return biased estimates of relationship parameters, and different methods in fact are not estimating the same parameter. Despite this variability in estimates, it is unclear if the relative change across environmental gradients is consistent across methodologies. Here, we simulate data sets across an hypothetical environmental gradient and estimate the size spectra parameter (slope or exponent, depending on method) at each site, and, importantly, estimate the relationship of parameters across the hypothetical gradient. We also use two previously published body size datasets across an anthropogenic stress gradient and an environmental temperature gradient and assess how the conclusions of these studies would vary based on methods used.
3. We find that maximum likelihood methods always perform better than common binning methods. Additionally, the variance in estimates using MLE methods is markedly reduced when compared to binning methods.
4. The uncertainty and variation in estimates when using binning methods is often greater than or equal to the variation previously published in experimental and observational studies, bringing into question the effect size of previously published results. However, re-analysis of two previously published datasets does not markedly alter the conclusions reached. Further study is needed to identify how and when these estimates may be biased, and when the general pattern of results can be considered to be “true.”

# Introduction

Body size distributions are a fundamental characteristic of communities. In general, size-abundance relationships are expected to scale with the value of -0.75 across all biological communities (Damuth 1981, 1991, 1998). This is thought to be a consequence of simple size-dependent metabolic constraints on organisms’ energy use predicted by the metabolic theory of ecology (Nee 1991, Brown et al. 2004). The remarkable consistency of these relationships across spatiotemporal scales and ecosystems has led them to be recommended as a “universal” indicator of ecological status (Petchey and Belgrano 2010). Variation in size-abundance relationships have been documented through space (Pomeranz et al. 2022), time (Evans et al. 2022), seasonality (McGarvey and Kirk 2018), to human activities (Jennings and Blanchard 2004). Likewise, variation in size spectra relationships have been used to explain fundamental differences in how communities are organized. For example, external resource subsidies “bend the rules” and allow higher abundances of large body sizes than would be expected (Perkins et al. 2018). However, recent research has shown that these results may be an artifact of how the data were treated. Edwards et al. (2020) reanalyzed a time series of marine fisheries data and found that previously reported changes through time were actually dependent on the methodology used.

Individual size distributions (ISD *sensu* (White et al. 2007), also referred to as abundance size spectra) are one of the size-abundance relationships commonly used. ISDs show the relationship between the abundances of individual body size regardless of taxonomy. Generally, there is a negative relationship between individual body size (M) on the *x*-axis and abundance (N) on the *y*-axis. Theoretical and empirical data support this relationship being described as a simple power law with exponent in the form of , where Andersen and Beyer (2006). Commonly, is the count of body sizes grouped into bins, and is estimated as the slope from OLS regressions of log transformed data of and the mid point of the body size bin : . Myriad binning methods have been proposed, including linear and logarithmic bin widths. Likewise, methods can rely on the absolute count in the bins as well as normalization techniques where the count is divided by the bin width (especially common with logarithmic binning). Alternatively, can be estimated directly using maximum likelihood techniques.

Previous work has shown that the estimates of differ between MLE and OLS techniques. OLS methods are particularly sensitive to decisions made in the binning process. Simulation studies have shown that MLE offers consistently more accurate estimates of (White et al. 2008, Edwards et al. 2017), and reanalysis of empirical data also indicates that the conclusions are dependent on the methodology used Edwards et al. (2020). However, recent empirical analysis of stream macroinvertebrate communities across the National Ecological Observatory Network (NEON, USA) showed that while the estimates of varied based on method, the relative change across the environmental gradient was consistent across methods (Pomeranz et al. 2022). While there is a growing consensus that MLE methods offer more reliable estimates of , and binning methods result in biased estimates, it remains unclear if these biases are consistent and systematic or stochastic, and whether or not the relative change in ISD parameters is consistent across space and time. In other words, if the data within a study are all treated the same, does a relative change of OLS slope parameters of 0.1 coincide with a relative change of MLE estimates of 0.1?

We had three primary objectives in this study: 1) how well do different methods estimate site-specific ’s? 2) Do the estimates of the hypothetical relationship obtained from different methods agree? and 3) Are the conclusions reached from previously published data sets dependent on the method used?

In order to answer the first two questions, we simulate body size observations from bounded power law distributions with varied exponents across a hypothetical environmental gradient and compare the results obtained using three different methodologies common in the literature. To answer the third question, we re-analyze two previously published data sets of community body sizes across a stress and environmental gradient, respectively Pomeranz et al. (2022). We find that the MLE method more accurately estimates the site-specific exponents as well as the relative change in exponents across the hypothetical gradient, as well as having smaller variation around the estimates. The logarithmic binning methods generally perform well, but have larger variation around the site-specific estimates. In addition, the variation of the estimates of the hypothetical relationship obtained from the binning methods is dependent on the value of used (under and over estimates steep and shallow , respectively).

# Methods

## Data Simulation

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

and . These values are based on empirical body sizes of stream benthic communities reported in (Pomeranz et al. 2022). Our results are not dependent on the range of body sizes, and we show the results of other body size ranges in the supplemental information.

## Site-specific estimates

We sampled body sizes from distributions described by five different ’s: (-1.5, -1.75, -2.00, -2.25, -2.5). The values of describe how quickly the abundance of large body sizes decline within a community, e.g., a value of -1.5 means there would be a relatively high number of large body sizes (shallow) whereas a value of -2.5 means there would be relatively very few large body sizes (steep). We used three different methods (described below) to estimate the value of . We repeated this process 1000 times (reps) and plot the distribution of values obtained for each method.

## Estimation of Size spectra parameter

For each sample of and each of the 1000 replicates, we estimated the exponent directly using MLE methods modified from the the sizeSpectra package (Edwards et al. 2020). Throughout the manuscript, these estimates are referred to as MLE.

### Equal Logarithmic Bins: ELBn

In addition, we use two common binning methods to estimate the OLS slope parameter in log-log space. It is important to note that the binning methods are estimating a slope () from an OLS regression analysis, and not directly. However, in the interest of clarity, we will refer to all estimates as regardless of the method used, and reserve the term to refer to the estimate of the size spectra relationships across the hypothetical environmental gradient.

For the first binning method, we created 6 equal logarithmic bins covering the range of body sizes. This method has been used extensively in previous studies. For example see Dossena et al. (2012). In the present study, the count in each bin was normalized by dividing by the bin width to account for the unequal bin sizes. The process of normalizing shifts the estimate by -1. In other words, an un-normalized estimate of -0.75 would result in an estimate of -1.75 when normalizing the data Pomeranz et al. (2022). Throughout the manuscript, the normalized equal logarithmic binning method will be referred to as ELBn.

### Log2 Bins: NAS

The second method was similar to ELBn, but bins of Log2 widths are used. The number of bins is dependent on the range of body sizes present in the data. When working with empirical data with different size ranges this can alter the number of bins per site, and it is recommended to construct the bins based on the global min and max body size present. However, since the data here were simulated from a known size range, the number of bins for each site is identical. The count in each bin is normalized in the same way as described above. This method has been used [Jennings and Blanchard (2004), Jennings et al. (2002), Sprules and Barth (2015), Pomeranz et al. (2019a); McGarvey and Kirk (2018)) and is referred to as the Normalized Abundance Spectrum (NAS).

For the two binning methods, simple OLS regression were conducted in the form

Where is considered to be the .

## Variation in Across a Gradient

### Hypothetical relationshiop of -0.5

Previous work has investigated the biases in estimates when using different methods White et al. (2007). However, the focus of the present work is to investigate biases when estimating the relationship parameters across an environmental gradient.

To do this, we again sampled body sizes from distributions described by five different ’s. However, because the estimates from the binning methods varied depending on the value of (see results), we chose three different “windows” of lambda values. The three windows were steep (), medium (), and shallow (().

**NOTE** This MS seems to be getting long. Possibly present “medium” window in main results, and include other windows in SI

Each sample of body sizes from different values was assumed to come from a community across a hypothetical environmental gradient . Values of for each community are uniformly distributed from -1 to 1 (site-specific x-values: . For each “window,” the relative change in values () across the gradient is equal to -0.5, such that our model is:

**JEFF** Is this correct? for some reason this seems wrong to me.

We repeated the data simulation 1000 times (reps). The main results presented here were not dependent on the range of *x*-values or the number of sites (supplemental information).

### Varying Hypothetical Relationship

Finally, we also varied the relationship () across the hypothetical gradient. We tested values of 0, -0.25, and -0.5 by sampling 5 body sizes communities uniformly distributed across the following sets of values: [-1.5], [-1.25, -1.75], [-1, -2].

## Estimation of relationship across gradient,

For each simulation replicate and size spectra method, we estimated how the size spectra parameters varied across the hypothetical gradient. We did this by performing simple OLS regression of the across the hypothetical gradient ():

The distribution of the relationship coefficient, , were plotted compared to the known relationship.

## Empirical Data

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

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

## Data Availabity

The data used in this manuscript is already publicly available (Pomeranz et al. 2019b) data dryad DOI <https://datadryad.org/stash/dataset/doi:10.5061/dryad.v6g985s>, (National Ecological Observatory Network (NEON) 2022). R Scripts to reproduce the full simulation and analysis are available at: ([<https://github.com/Jpomz/detecting-spectra-differences>]). (To be archived upon acceptance.)

# Results

# estimates

There was considerable variation in the estimate across methodologies (Figure 1). The distribution of estimates from the MLE method was almost always symmetrical and centered at the known value of (Figure 1, blue distribution and dashed line). The distribution of estimates from the binning methods were generally wider, and sometimes showed assymetrical distributions.

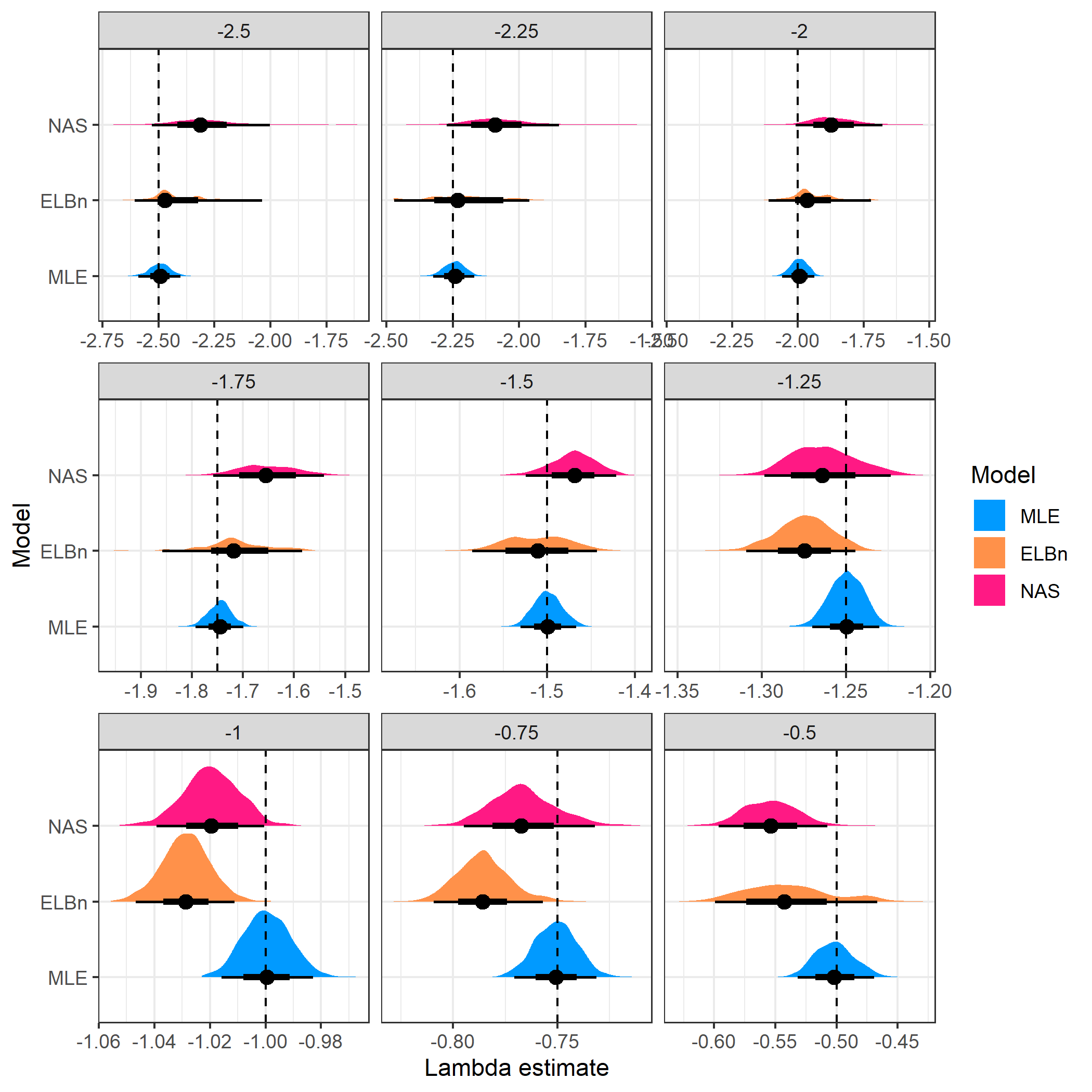


Figure 1. Distribution of Lambda estimates by method (color) from random samples of body sizes from bounded power law distributions with varying exponents (-2.5, -0.5). The figure is facetted by the known lambda parameter (facet title) and is also shown as the dashed line in each facet. Note that the x-axis varies in each facet.

Furthermore, the binning methods systematically over estimated steep size spectra relationships (, Figure 2), and slightly underestimated shallow size spectra relationships (). This finding was more pronounced in the NAS method compared with the ELBn method.

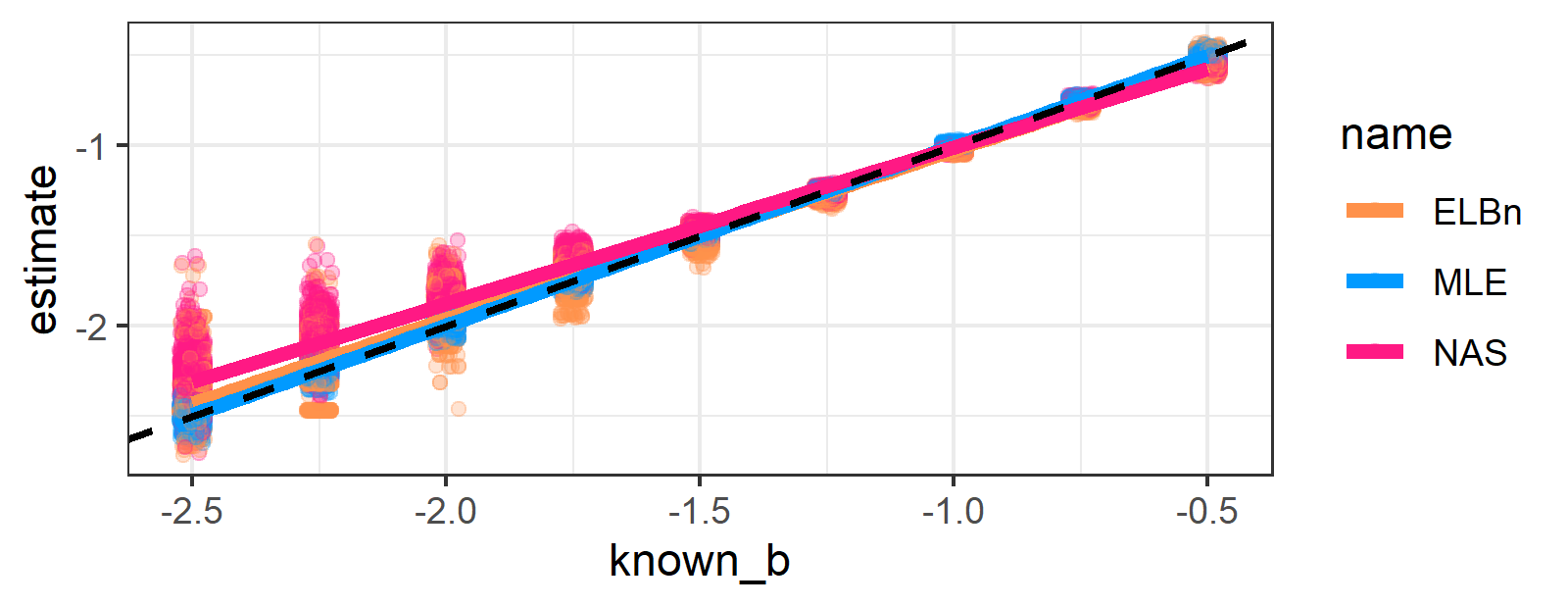


Figure 2. Plot of the known and estimated lambda from different methods (color). The dashed line is the 1:1 line, and almost perfectly matches the MLE (Blue) estimate. The two binning method cross the 1:1 line, meaning they systematically underestimate shallow lambdas, and over estimate steeper lambdas.

## Relationship across the hypothetical environmental gradient

### “windows”

Because of the different performance of the two binning methods at steep and shallow values of lambdas, we first investigated how well the methods capture a known relationship of across a hypothetical environmental gradient. The lambda “windows” are steep (-2.5, -1.5), medium (-2, -1) and shallow (-1.5, -0.5).

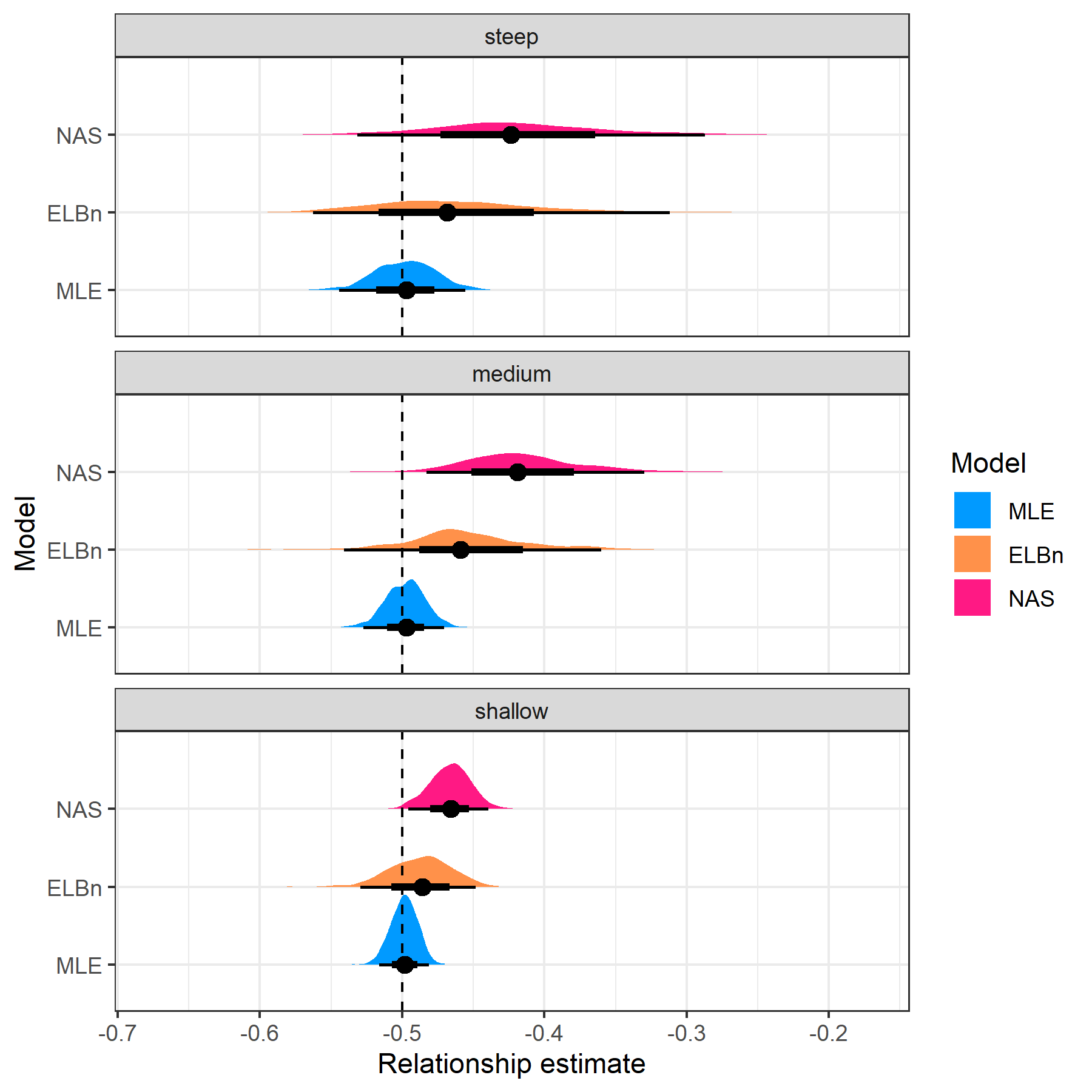


Figure 3. Distribution of relationship estimates () in three different “windows” of lambda values. The dashed vertical line is the known relationship value of -0.5.

The distribution of relationship estimates from the MLE method (Figure 3, blue) is approximately centered at the known relationship value in each of the “windows.” The distribution for the MLE estimate is widest in the “steep” window, but is still markedly narrower than the binning methods. In addition to having much wider distributions, the binning methods systematically overestimate the known relationship.

### Varying the Known Relationship

Finally, we were interested in finding if the three methods performed differently with varying values of the known relationship. Here, we present estimates from each method when the known relationship () across the hypothetical gradient was -0.5, -0.25, and 0, respectively (Figure 4).

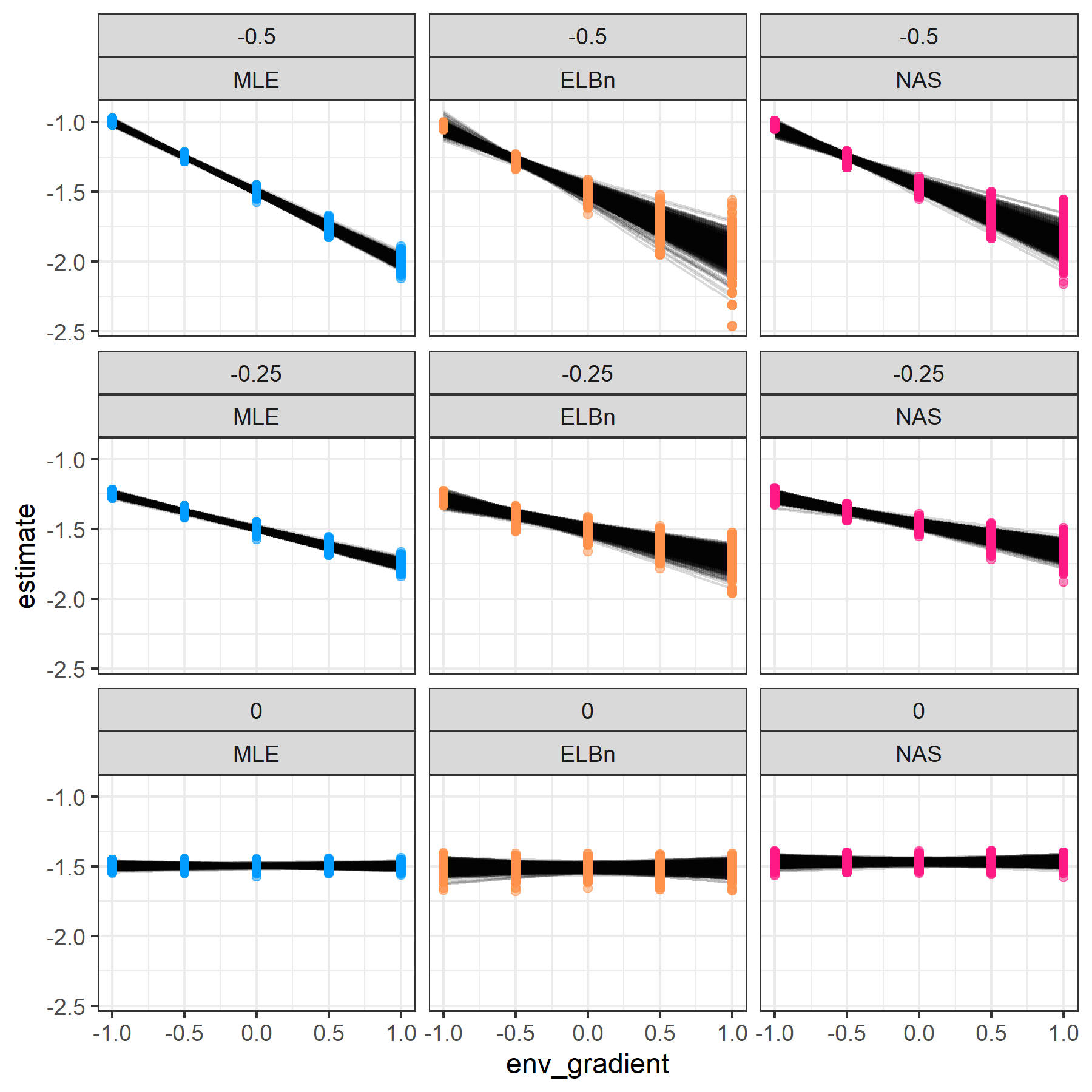


Figure 4. Individual regressions for each rep (N = 1000) and method (columns) for three different known relationship values (rows): 0, 0.25, and 0.5

Likewise, we also present the distribution of the relationship estimates (Figure 5). All the methods perform similarly and reasonably well when no relationship is present. The performance of the binning methods declines (center of distributions), with relationships being systematically over estimated. Likewise, the width of the distributions increases with stronger relationships across a hypothetical gradient. However, the distribution of estimates from the MLE method is nearly always centered at the known value, and has relatively narrow distributions.

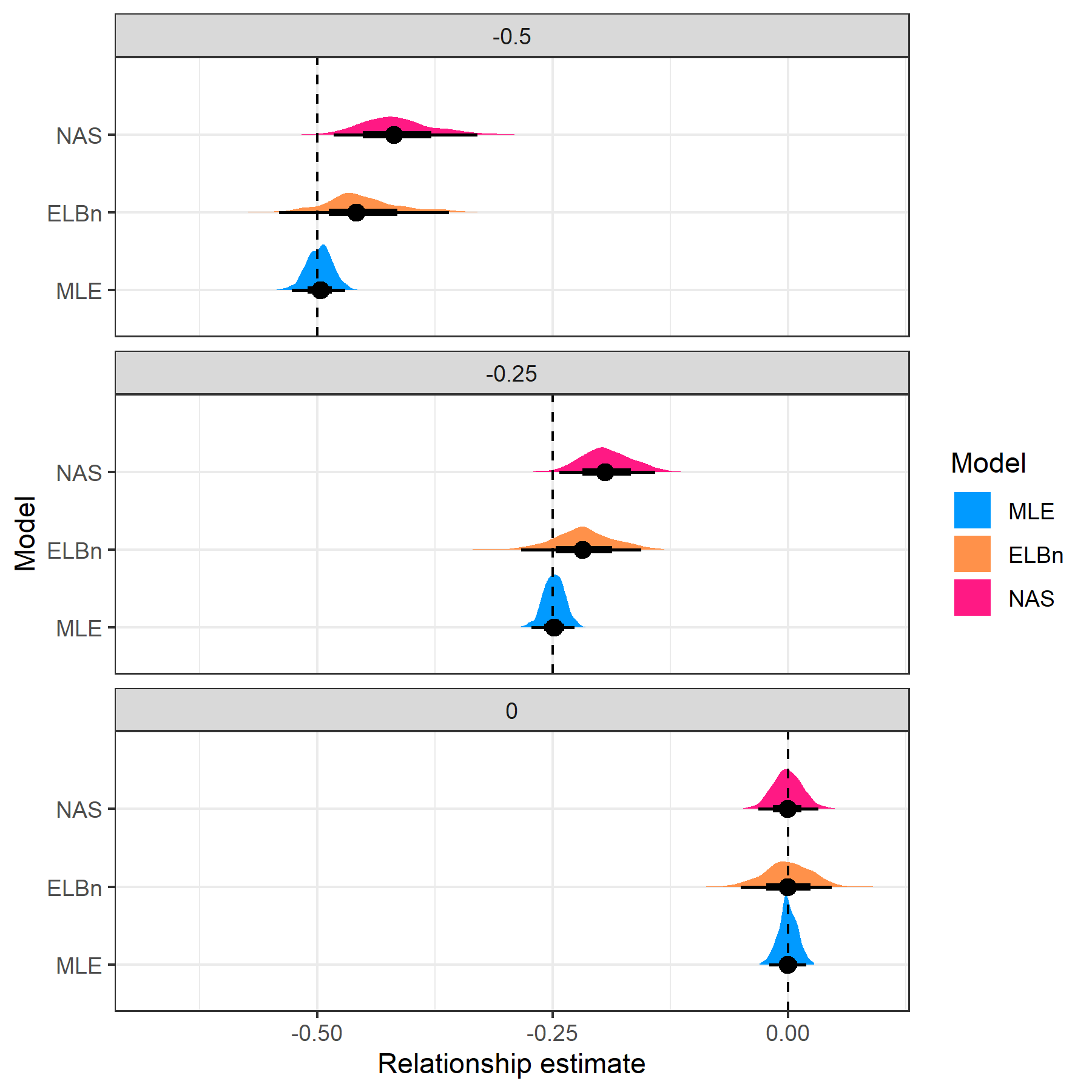


Figure 5. Distribution of relationship estimates (beta\_1) when estimating from different known relationships

## Empirical data

For both empirical data sets, the direction and magnitude of change (i.e.  coefficients) are generally in agreement. Size spectra parameters consistently increase (become flatter) in the AMD data (Fig. 6A). However, the magnitude of the change across the AMD gradient is weakest using the MLE method, and the width of the estimates ( 1 SD) is larger in the binning methods, particularly the ELBn method (Fig. 6B). Likewise, the size spectra parameters consistently increase (become steeper) with increasing temperature across the NEON sites (Fig. 6C). The magnitude of the relationships is weakest in the MLE method, although the uncertainty around the estimates are similar (Fig. 6D).

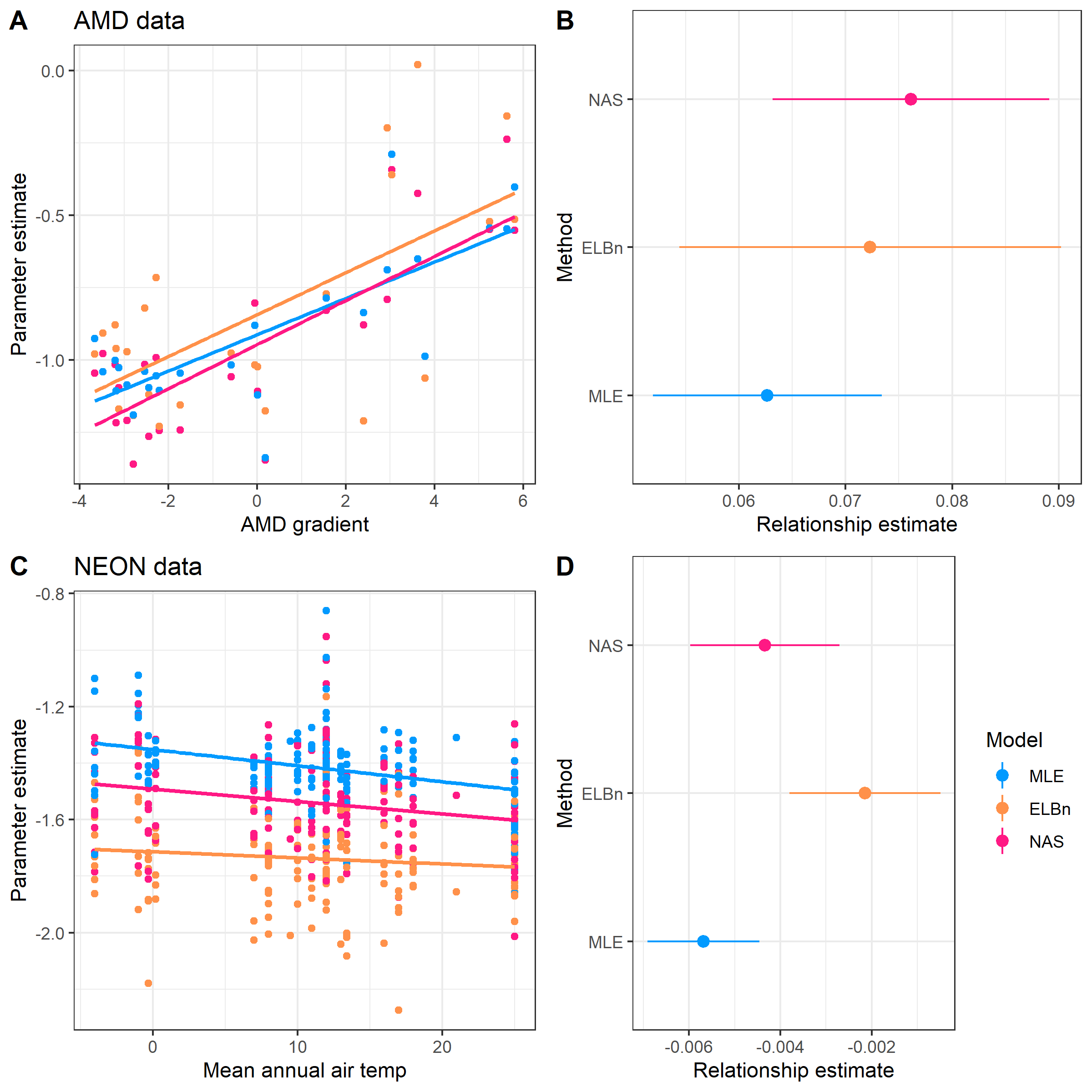


Figure 6. of empirical data estimates. All of the methods estimate the same sign of the relationship, but the estimates from the binning methods are generally greater than the MLE estimates.

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

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

# Discussion

The relationship between body size and abundance has been extensively studied in a wide range of taxa inhabiting both terrestrial and aquatic ecosystems (reviewed by White et al. (2007)). Empirical patterns are remarkably consistent, and can be explained by the metabolic theory of Ecology (Brown et al. 2004). Measuring parameters describing the decline in abundance with increasing body size in communities is being done with increasing frequency across ecology. Previous work has investigated the accuracy and inherent biases associated with different estimation methods. However, how these inaccuracies and biases compound across environmental gradients remains uncertain, making it difficult to detect variation in size spectra parameters across environmental gradients with confidence. Here, we sampled body sizes from known distributions with varying parameters () and estimated the coefficient of the relationship () across a hypothetical gradient in order to assess how the results would vary depending on the method used. Likewise, we compared how the interpretation of previously published results could change depending on the methodology used to estimate the relationship.

Generally, the estimate of the size spectra parameter from the MLE method was always closer to the known value, regardless of how steep or shallow the exponent was. This is in agreement with Edwards et at. (2017) which found that the MLE estimate more consistently recovered known values ranging from -2.5 to -0.5. We further this work by discovering that the magnitude of the error associated with the two binning methods increases as the value moves away from -1.

It is unsurprising that all methods performed less well in estimating steep values of (i.e., -2.5). The rapid decline of large body sizes makes it unlikely to sample any individuals with a large body size. Furthermore, the binning methods inherently “lose” data by aggregating individual observations into a single bin.

Likewise, the MLE method also resulted in the most accurate estimates for the hypothetical relationship in simulated data regardless of the “window” investigated. Furthermore, the variation around the estimate was always smallest with the MLE method. However, the variation in the MLE estimate did increase with steeper “windows.” This is not surprising, however because steep relationships like this make it very unlikely to sample the largest body sizes, and capturing the “true” relationships without the full range of body sizes present is more difficult (Edwards et al. 2017). Despite this, the MLE method is nearly always the best option to use, likely because all the data are used in the analysis.

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

The variation in estimated relationships across the empirical datasets varied by ~0.1 units depending on the method used. At first glance this may not appear to be a concerning magnitude of variation. However, when one considers the range of variation reported in many seminal works on variation in size spectra relationships is ~0.1-0.2 (SI Table *xx*), it becomes clear how important an error in estimates of this magnitude is.

Despite the differences in magnitude of the empirical relationship coefficients, they were generally in the correct direction and of a similar magnitude. This suggests that previously reported significant changes in size spectra parameters across environmental gradients and in experimental manipulations are plausible. Given that all of the data within a study is treated identically, the over all change in size spectra parameters is likely reasonable. However, the biases and inconsistencies in relationship estimates presented here suggest that it would be difficult if not impossible to directly compare the relative changes across different published studies which use different methods. However, the relatively systematic error of estimates across known exponents suggests that it may be possible to correct previous estimates onto the same scale (Figure 2).

Despite this possibility, publication of individual body size data with future studies of size spectra relationships would greatly aid in our ability to generalize changes to this fundamental aspect of community organization across spatiotemporal scales and in response to environmental conditions.

## Concluding Remarks

The MLE method outperformed both of the binning methods under nearly any measure. With the publication of the *sizespectra* package (2017), producing MLE estimates of size spectra parameters is a relatively easy task. Therefore, we recommend using this in all future studies of size spectra relationships.

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

# References

Andersen, K. H., and J. E. Beyer. 2006. Asymptotic Size Determines Species Abundance in the Marine Size Spectrum. The American Naturalist 168:54–61.

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago, IL.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771–1789.

Dossena, M., G. Yvon-Durocher, J. Grey, J. M. Montoya, D. M. Perkins, M. Trimmer, and G. Woodward. 2012. Warming alters community size structure and ecosystem functioning. Proceedings of the Royal Society B 279:3011–3019.

Edwards, A. M., J. Robinson, J. Blanchard, J. Baum, and M. Plank. 2020. Accounting for the bin structure of data removes bias when fitting size spectra. Marine Ecology Progress Series 636:19–33.

Edwards, A. M., J. Robinson, M. Plank, J. Baum, and J. Blanchard. 2017. Testing and recommending methods for fitting size spectra to data. Methods in Ecology and Evolution 8:57–67.

Evans, T. M., Z. S. Feiner, L. G. Rudstam, D. M. Mason, J. M. Watkins, E. D. Reavie, A. E. Scofield, L. E. Burlakova, A. Y. Karatayev, and W. G. Sprules. 2022. Size spectra analysis of a decade of Laurentian Great Lakes data. Canadian Journal of Fisheries and Aquatic Sciences 79:183–194.

Jennings, S., and J. L. Blanchard. 2004. Fish abundance with no fishing: Predictions based on macroecological theory. Journal of Animal Ecology 73:632–642.

Jennings, S., K. J. Warr, and S. Mackinson. 2002. Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. Marine Ecology Progress Series 240:11–20.

Martínez, A., A. Larrañaga, A. Miguélez, G. Yvon-Durocher, and J. Pozo. 2016. Land use change affects macroinvertebrate community size spectrum in streams: The case of Pinus radiata plantations. Freshwater Biology 61:69–79.

McGarvey, D. J., and A. J. Kirk. 2018. Seasonal comparison of community-level size-spectra in southern coalfield streams of West Virginia (USA). Hydrobiologia 809:65–77.

National Ecological Observatory Network (NEON). 2022. Macroinvertebrate collection (DP1.20120.001). National Ecological Observatory Network (NEON).

Perkins, D. M., I. Durance, F. K. Edwards, J. Grey, A. G. Hildrew, M. Jackson, J. I. Jones, R. B. Lauridsen, K. Layer-Dobra, M. S. A. Thompson, and G. Woodward. 2018. Bending the rules: Exploitation of allochthonous resources by a top-predator modifies size-abundance scaling in stream food webs. Ecology Letters.

Petchey, O. L., and A. Belgrano. 2010. Body-size distributions and size-spectra: Universal indicators of ecological status? Biology Letters 6:434–437.

Pomeranz, J. P. F., J. R. Junker, and J. S. Wesner. 2022. Individual size distributions across North American streams vary with local temperature. Global Change Biology 28:848–858.

Pomeranz, J. P. F., H. J. Warburton, and J. S. Harding. 2019a. Anthropogenic mining alters macroinvertebrate size spectra in streams. Freshwater Biology 64:81–92.

Pomeranz, J. P. F., H. J. Warburton, and J. S. Harding. 2019b. Data from: Anthropogenic mining alters macroinvertebrate size spectra in streams. Dryad.

Sheldon, R. W., and S. R. Kerr. 1972. The Population Density of Monsters in Loch Ness1. Limnology and Oceanography 17:796–798.

Sprules, and Barth. 2015. Surfing the biomass size spectrum: Some remarks on history, theory, and application. Canadian Journal of Fisheries and Aquatic Sciences 73:477–495.

White, E. P., B. J. Enquist, and J. L. Green. 2008. On estimating the exponent of power-law frequency distributions. Ecology 89:905–912.

White, E. P., S. K. M. Ernest, A. J. Kerkhoff, and B. J. Enquist. 2007. Relationships between body size and abundance in ecology. Trends in Ecology & Evolution 22:323–330.

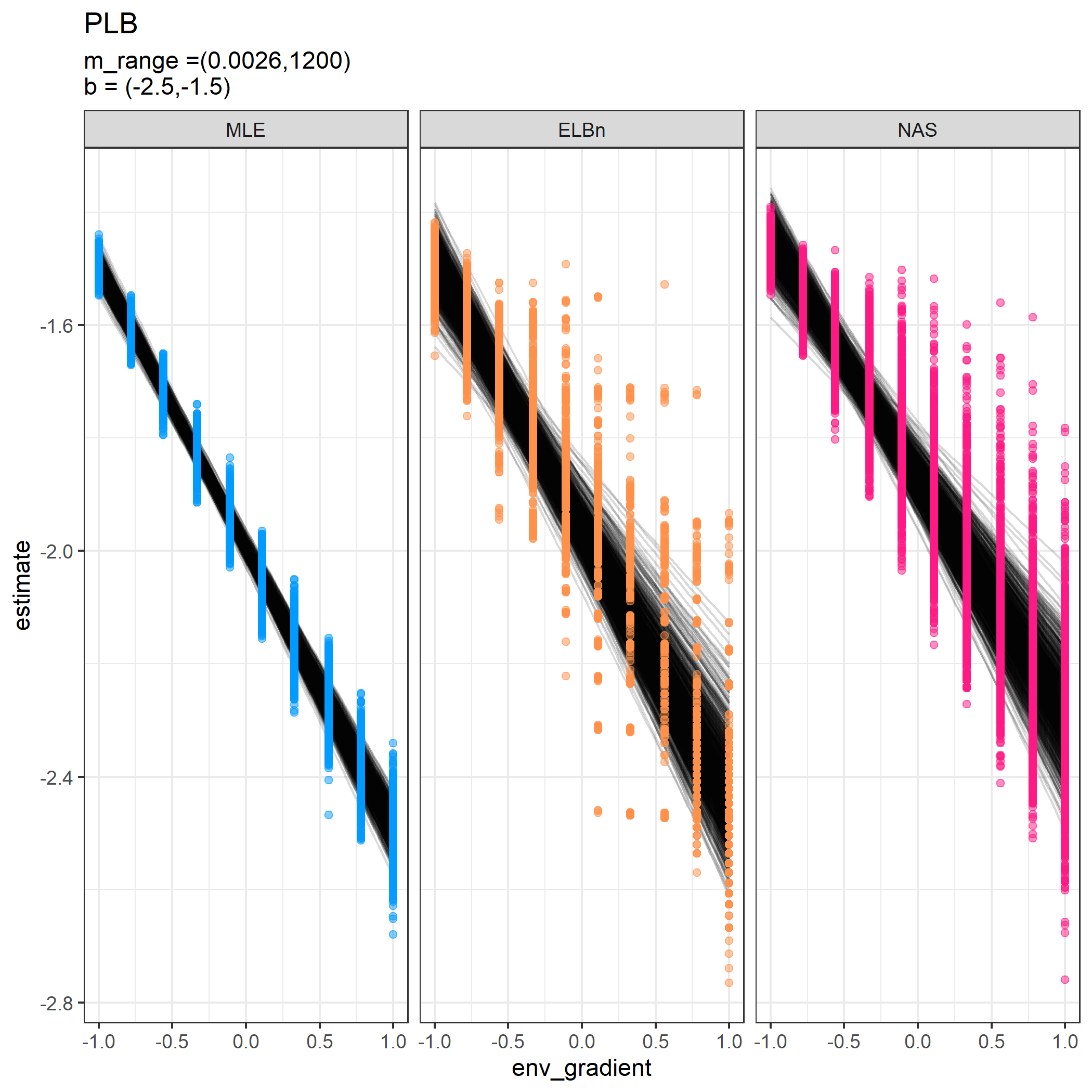
# Supplementary material

# and relationship estimates

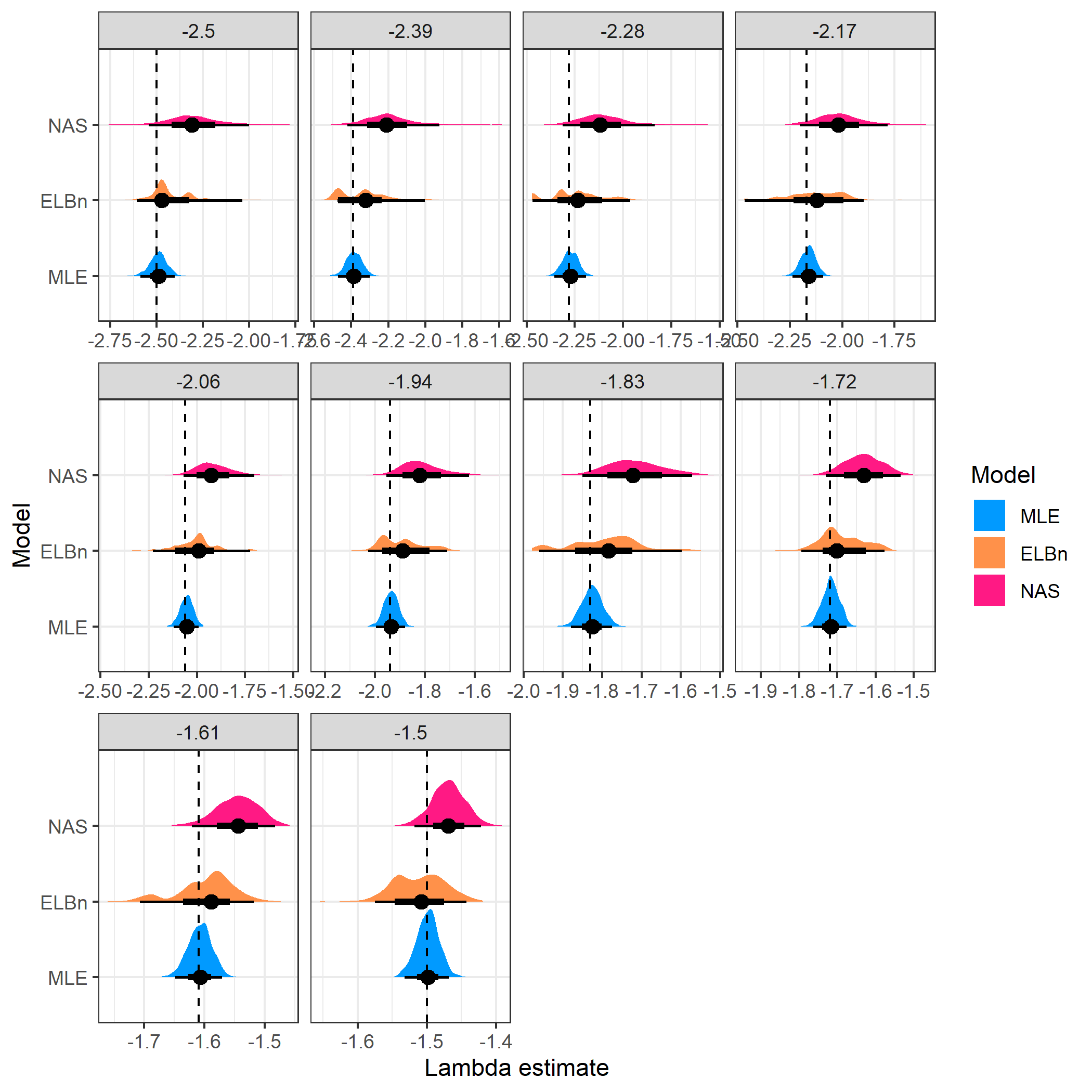
In the main analysis, we simulate body size data from bounded power law distributions while varying the exponent which describes the distribution. For the results presented in the main text, we held the number of sites at 5, scaled the environmental gradient from -1 to 1, and set the minimum and maximum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

## Varying number of sites

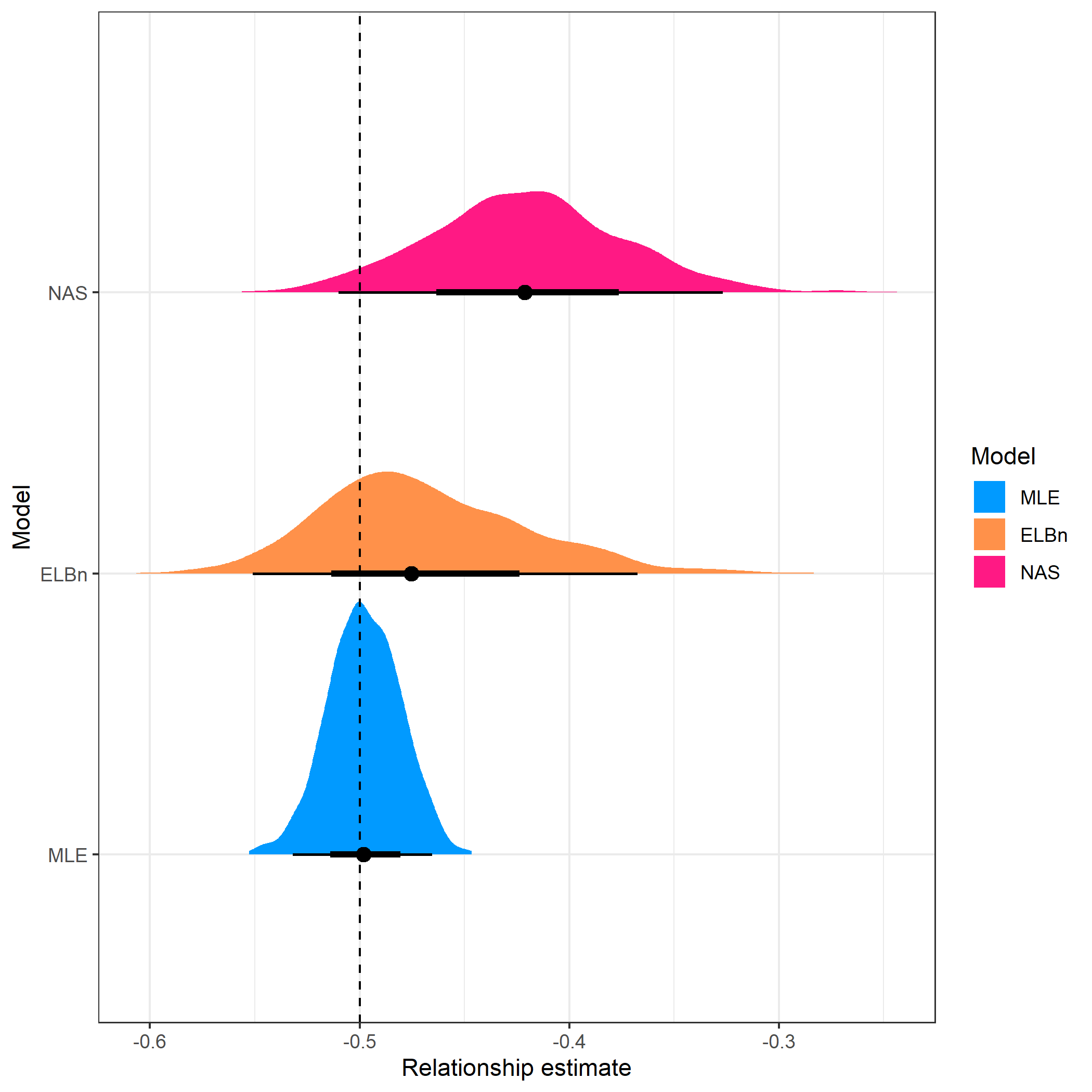
### 10 sites



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

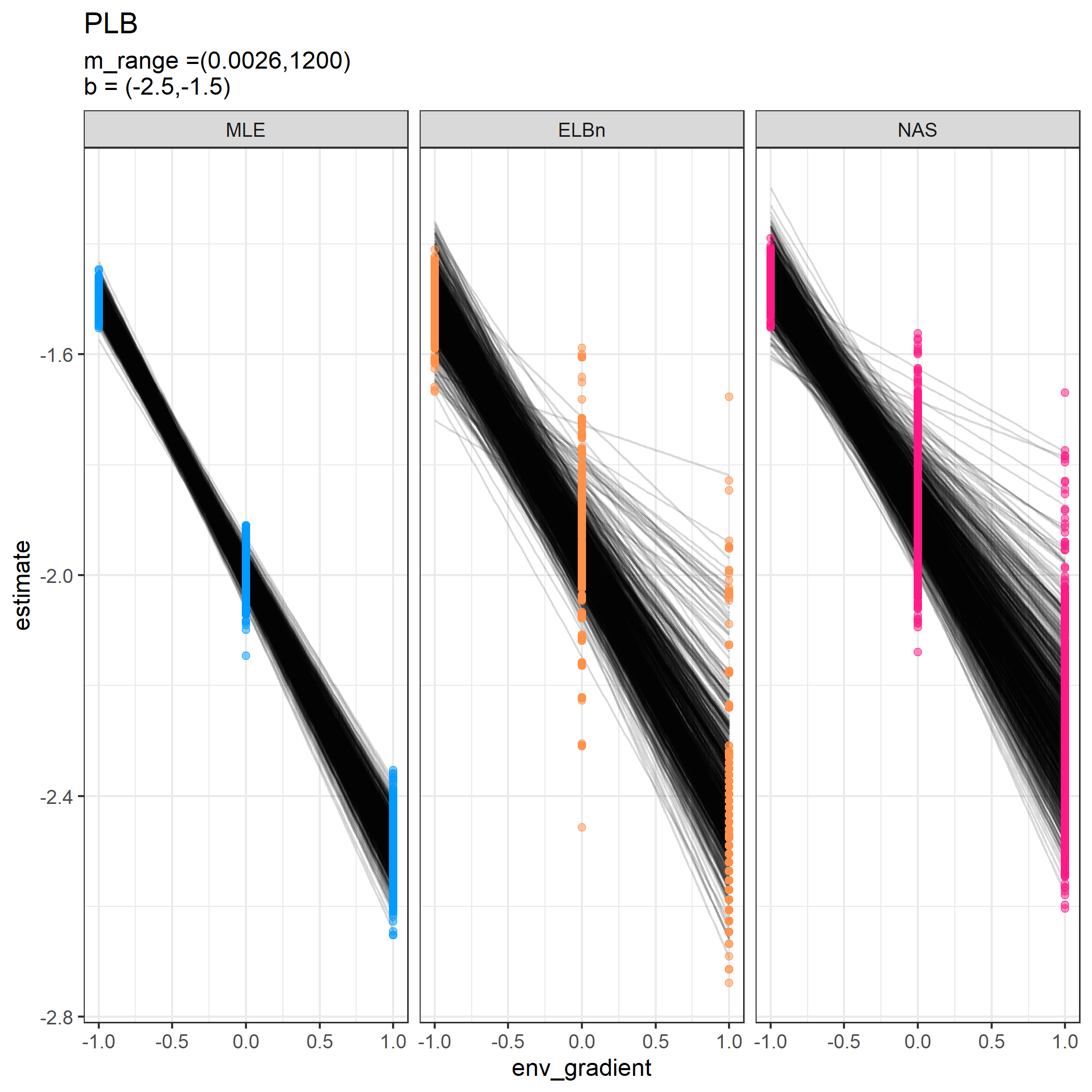


Distribution of estimated coefficient for ten sites across a hypothetical gradient with known values. All other parameters are the same as in the main analysis

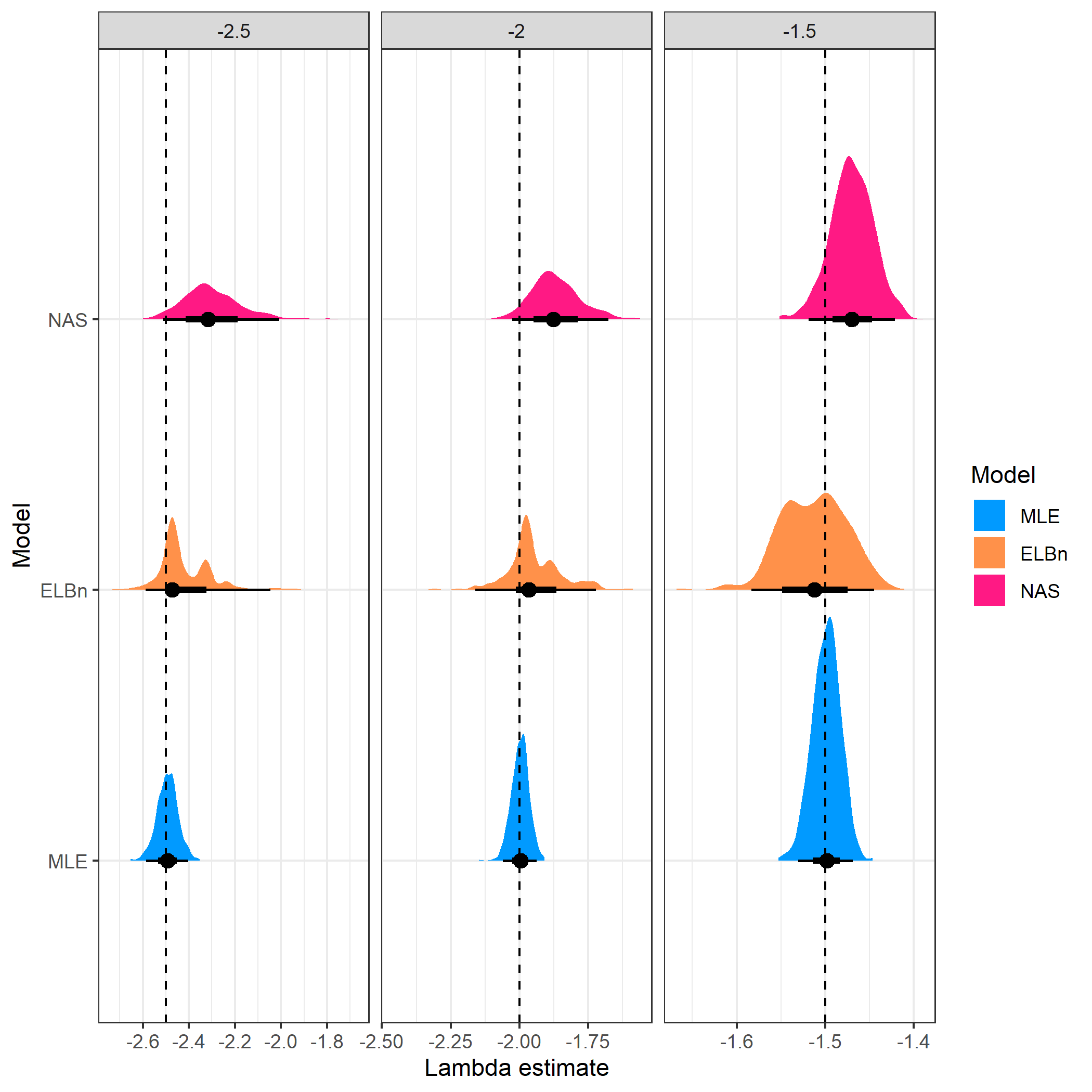


Distribution of estimated relationship () coefficient’s for ten sites across a hypothetical gradient with known value of 0.5. All other parameters are the same as in the main analysis

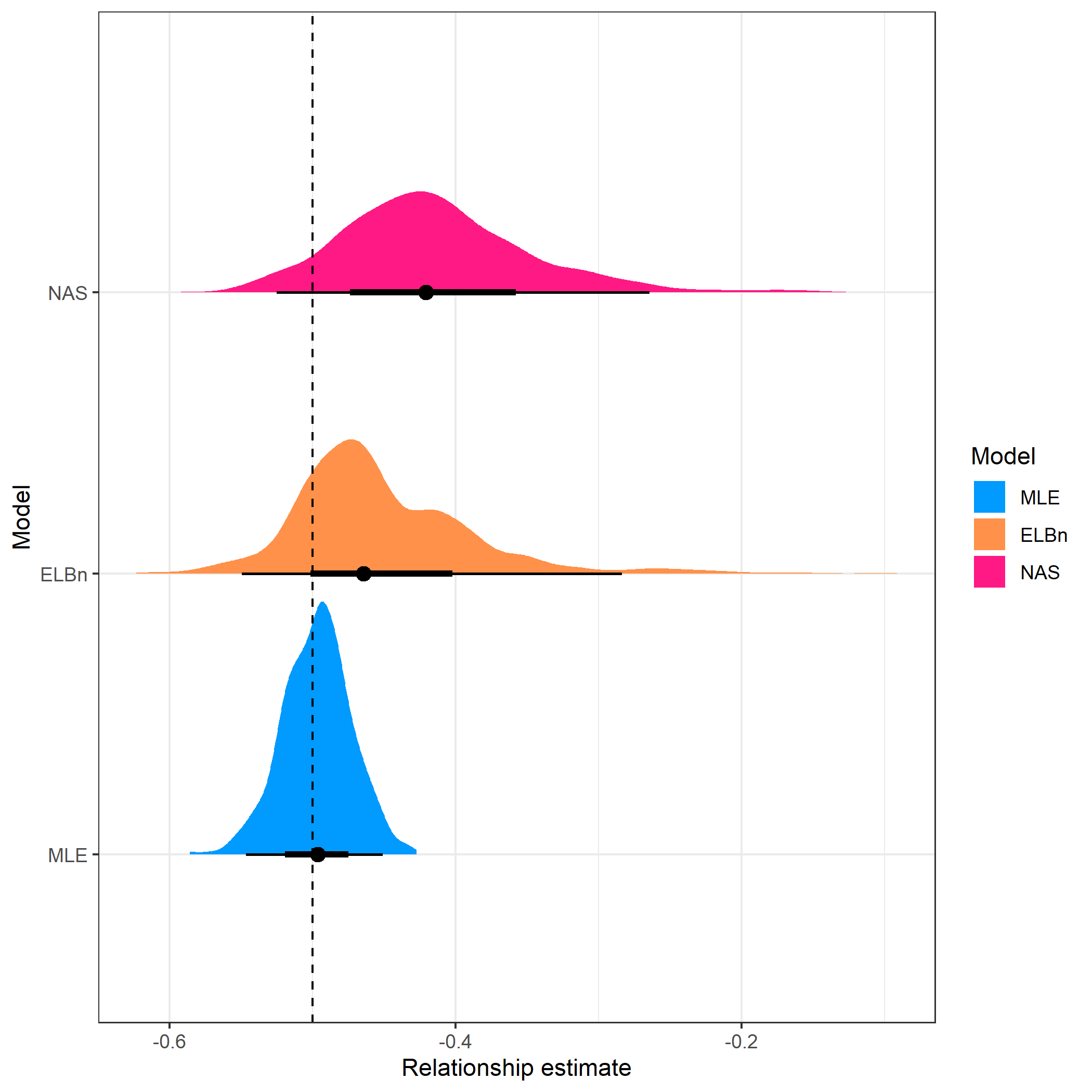
### Three sites



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.

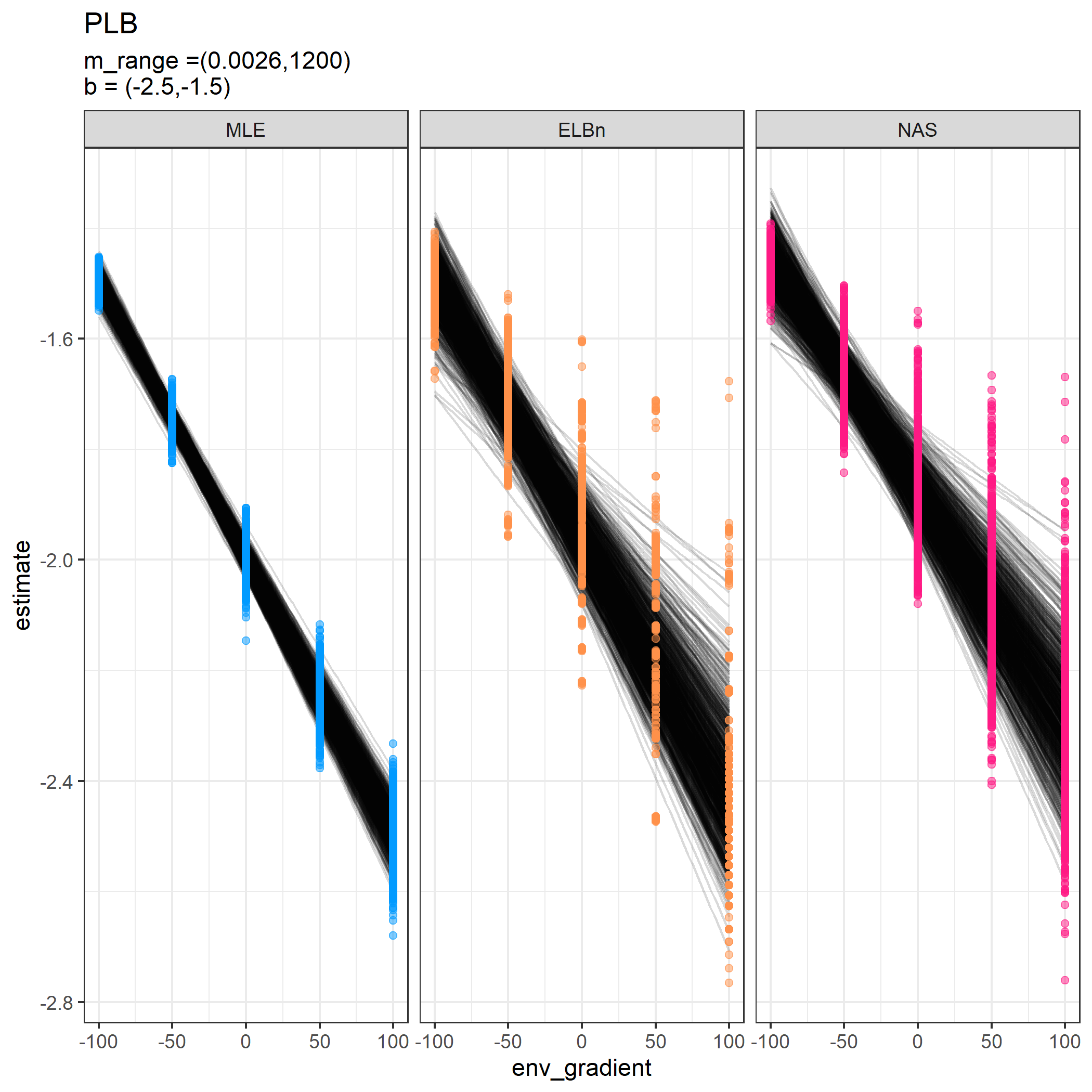


Distribution of estimated coefficient for three sites across a hypothetical gradient with known values. All other parameters are the same as in the main analysis.

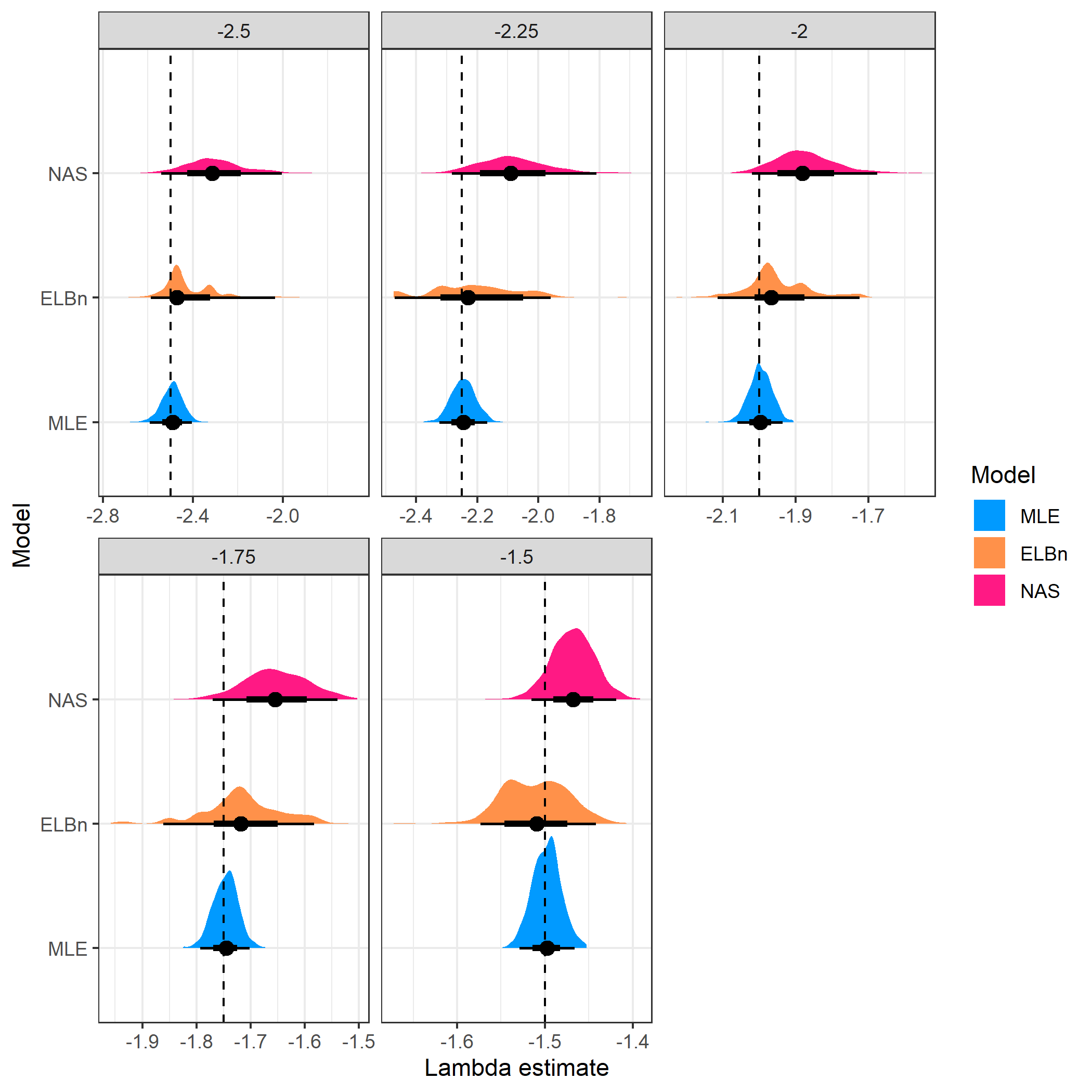


Distribution of estimated relationship () coefficient’s for three sites across a hypothetical gradient with known value of 0.5. All other parameters are the same as in the main analysis

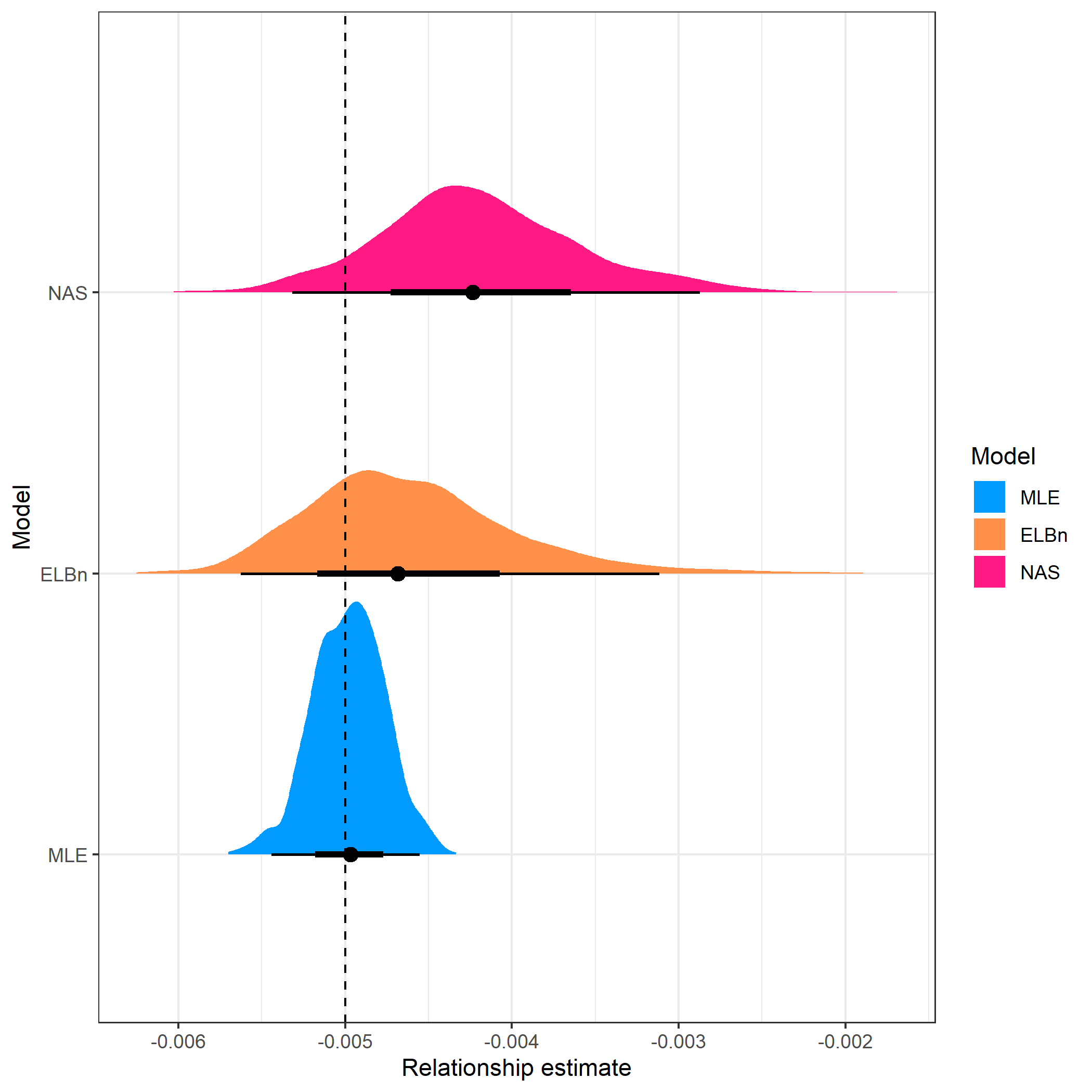
## Large environmental gradient



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.

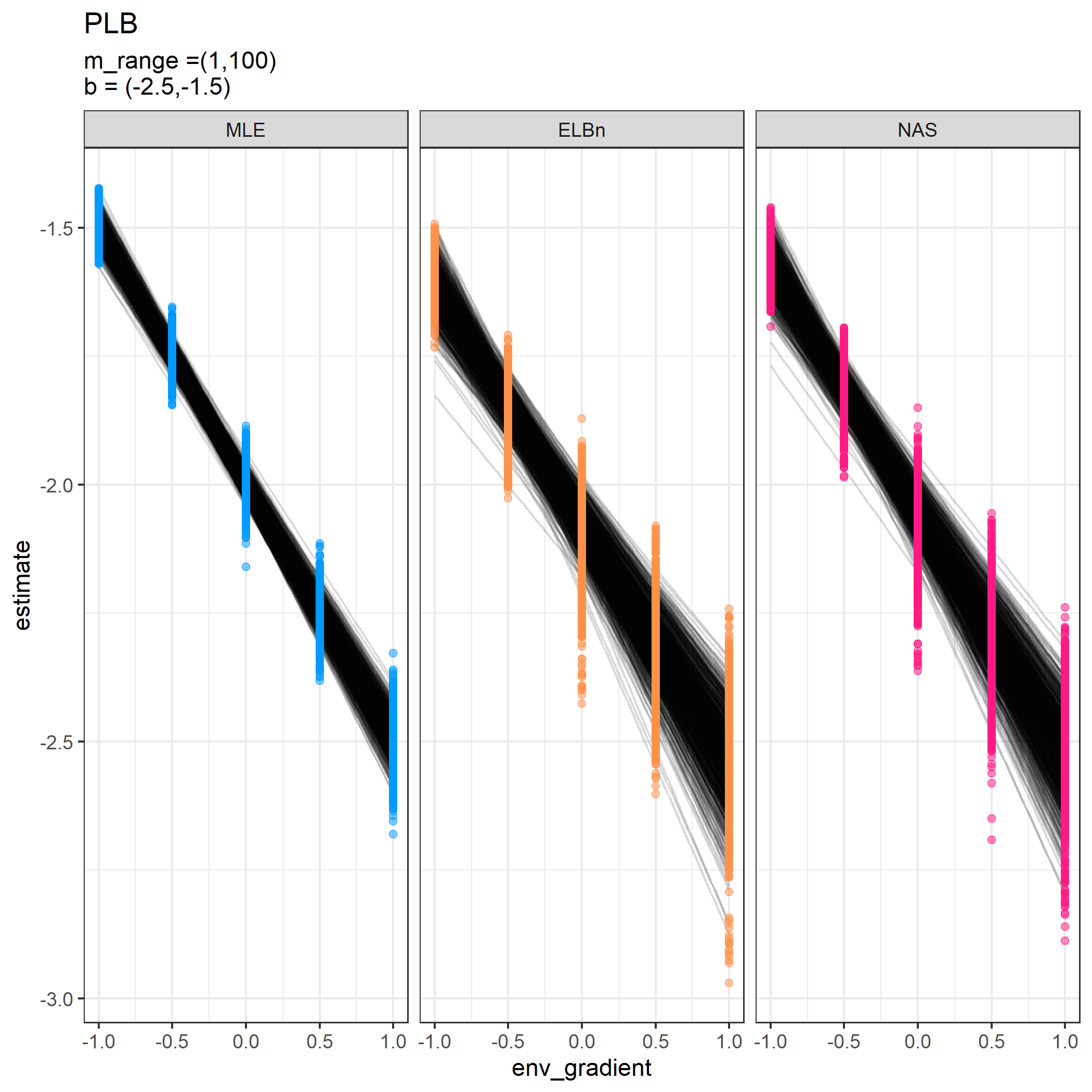


Distribution of estimated coefficient for five sites across a hypothetical gradient with known values. Range of environmental values (*x*-axis) increased to be -1000, to 1000.

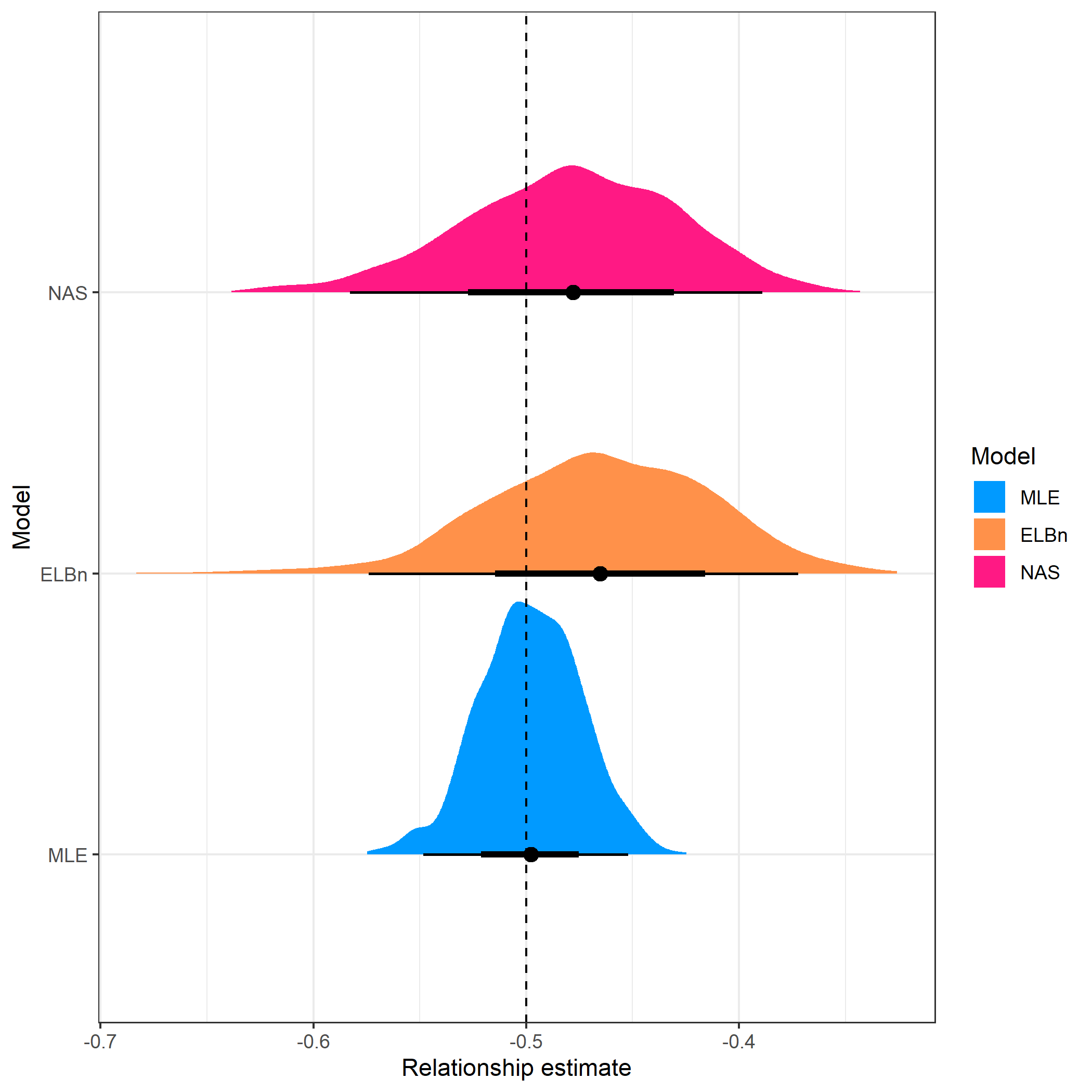


Distribution of estimated relationship () coefficient’s for five sites across a hypothetical gradient with known value of 0.5. Range of environmental values (*x*-axis) increased to be -1000, to 1000.

## Range of body sizes,



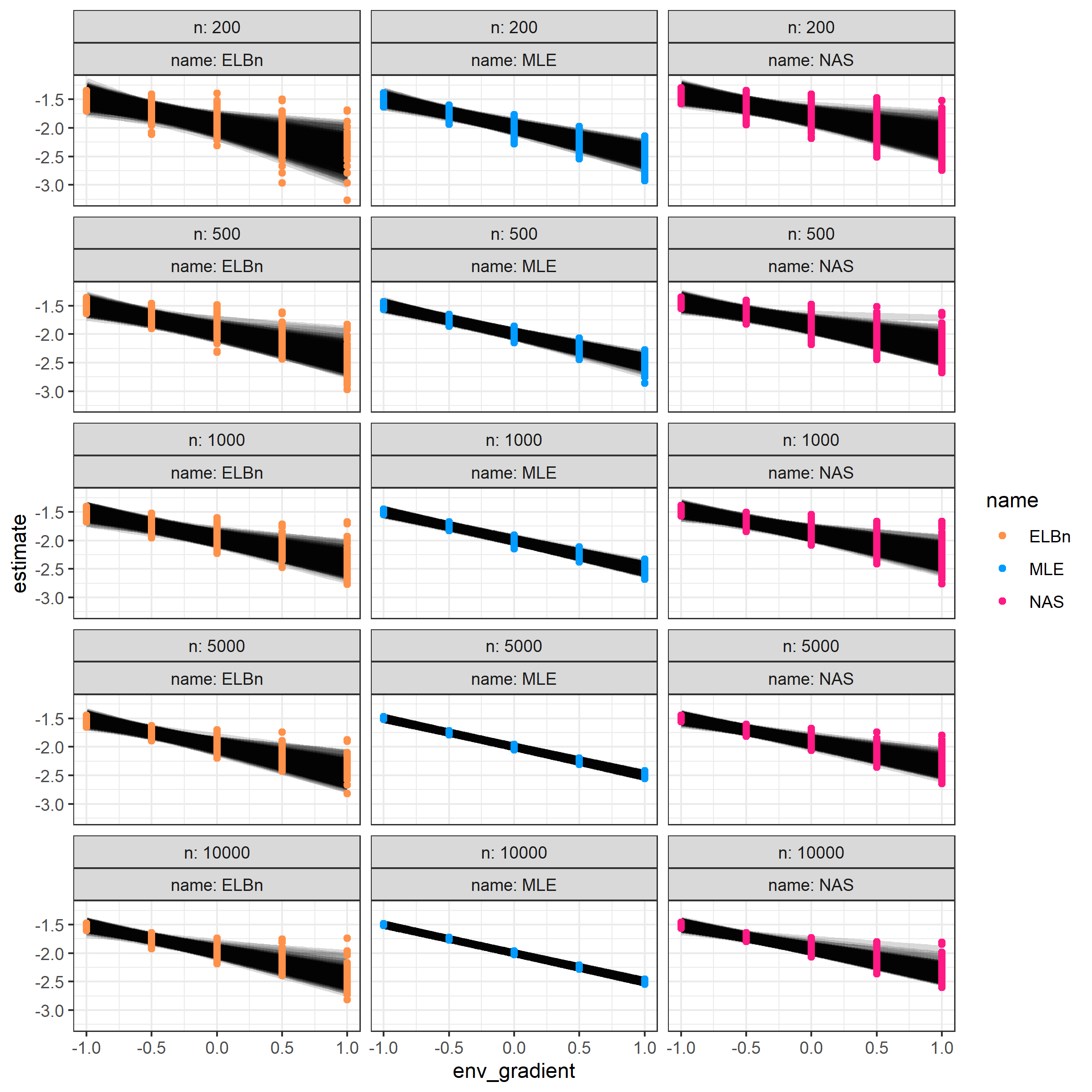
Individual regressions for five sites across a hypothetical gradient with a known relationship of 0.5. Range of body sizes is reduced and is from 1, to 100.



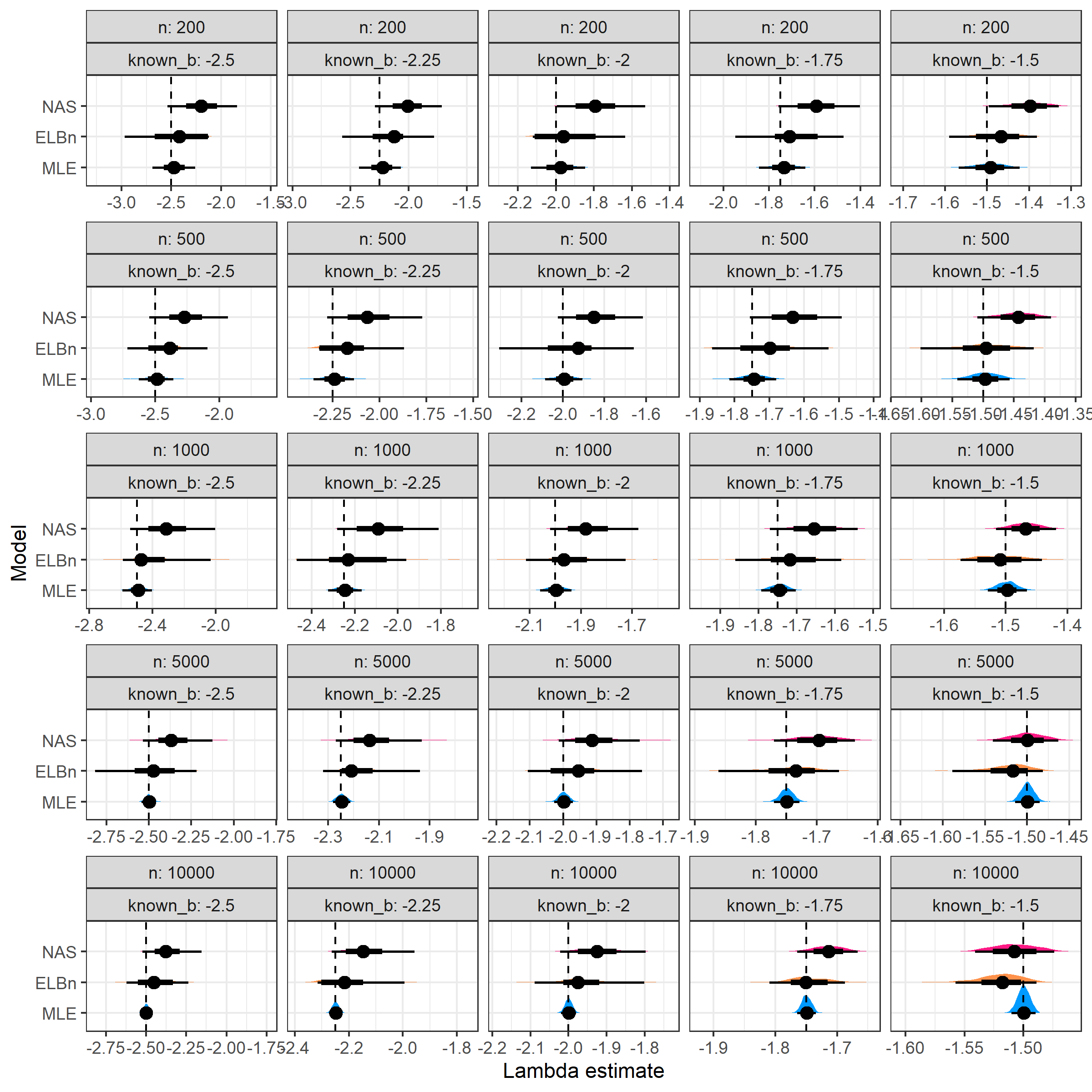
Distribution of estimated relationship () coefficient’s for five sites across a hypothetical gradient with known value of 0.5. Range of body sizes is reduced and is from 1, to 100.

## Sample size,

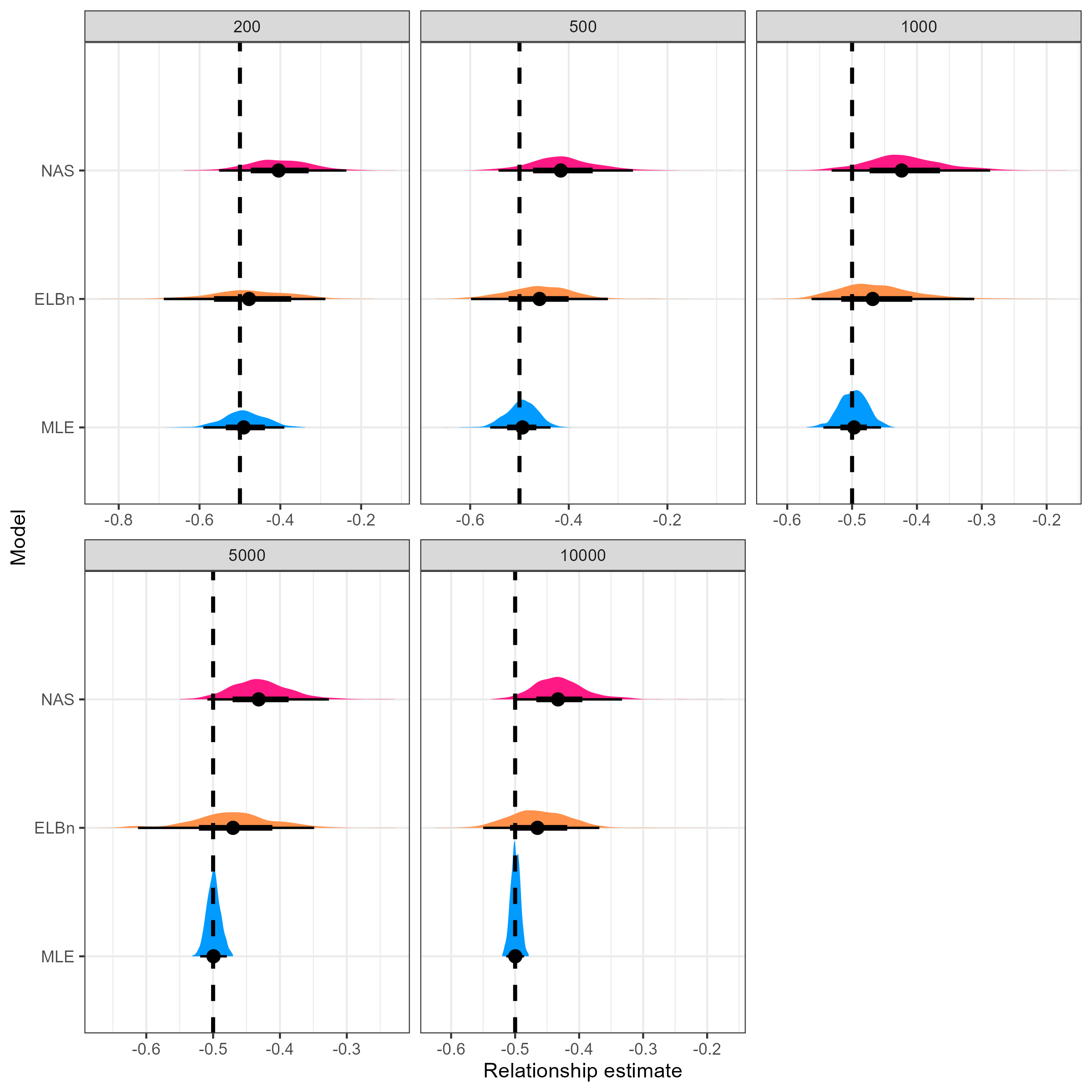
The number of observations in our simulations may bias the results. Therefore, we repeated the simulations described above, but varied the sample size . We tested values of . Results of this analysis are presented in the Supplemental Information.



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)

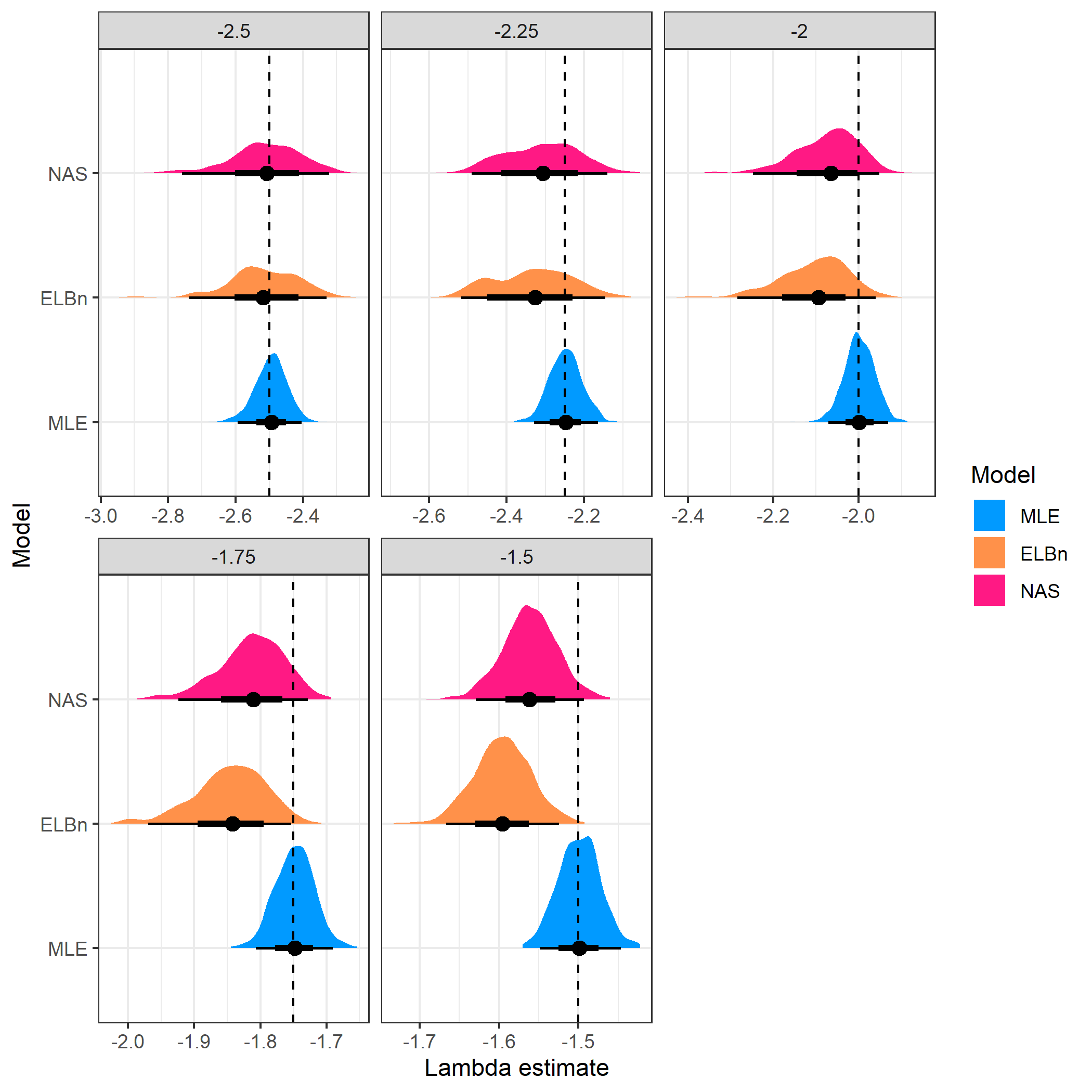


Distribution of size spectra parameter estimates. Vertical line is the known parameter (dashed line) wich describes the bounded power law distribution from which the body size estimates were sampled. As n increases (top to bottom) and increases (left to right), the accuracy of the estimate improves across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

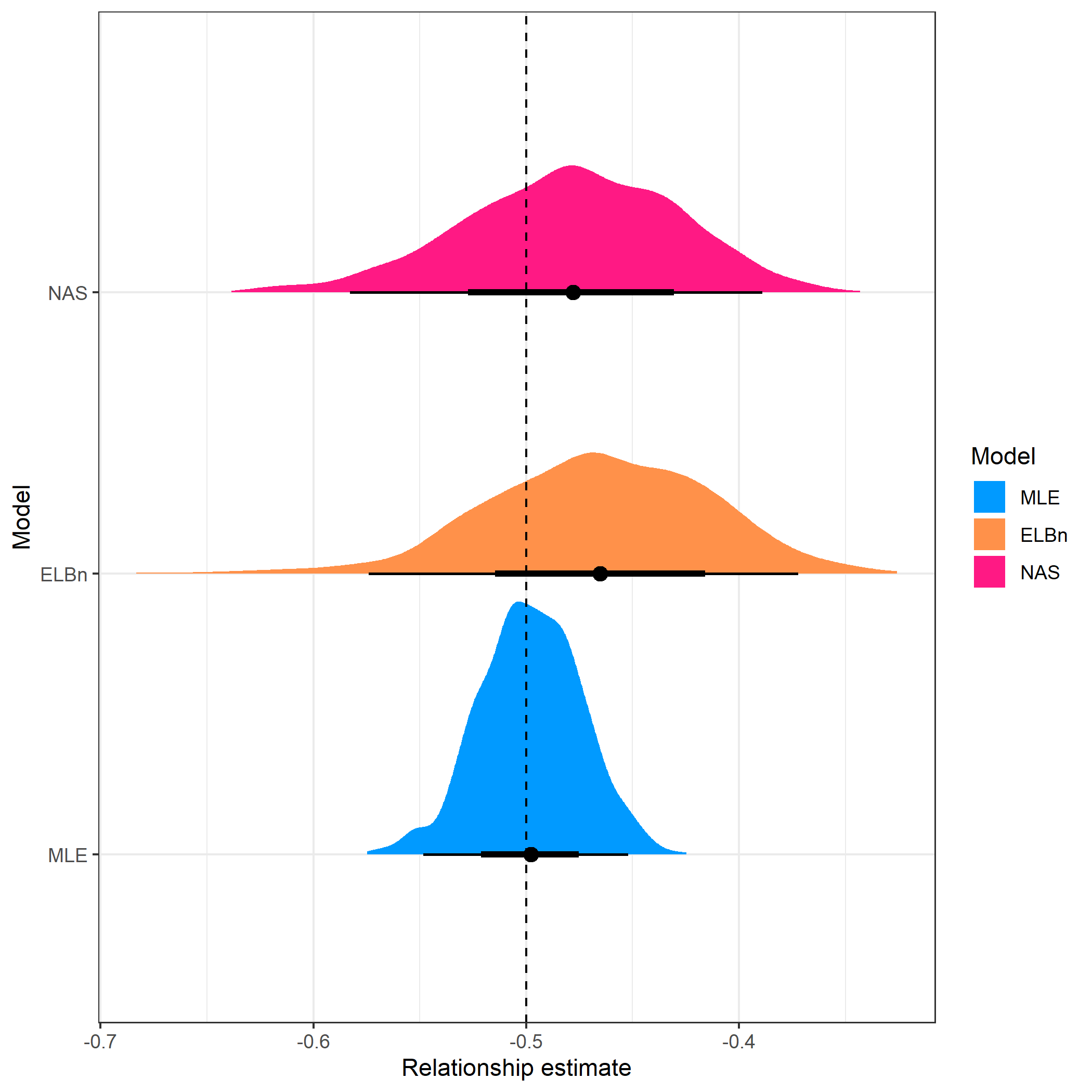


Distribtuion of the relationship coefficients with varying sample size

### Range of body sizes = 1 to 100



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 anaysis and ranges from 1, to 100.



Relationship estimate

# Comparison with other published estimates

SI Table. This table shows published estimates of the variation in size spectra slopes (or exponents) in empirical studies. It is unclear how to directly compare estimates of the slope with different methods.However, the published estimates here range from ~0.1 to 0.2 across the gradients studied. For comparison, the 2.5-95% quantiles around the relationship estimate for the MLE method were ~0.1, whereas for the ELBn and NAS method they were ~0.25 and ~0.2, respectively. b\_diff is the change in estimate (b-low - b-min). System refers to stream communities or mesocosm experiments. Method: MLE = maximum likelihod estimate, ELB = equal logarthmic binning, the number before indicates the number of bins used. The normalization process shifts the estimates by an absolute value of 1.0. Hence, direct comparison the relative change in normalized and non-normalized studies should not introduce any bias. The O’Gorman et al. 2017 study used average species size and abundance (Local Size Density Relationship, *sensu* White et al. 2007) as opposed to individual size distribution. These methods are related, but it is unclear how to directly compare estimates from each method.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Citation | Difference | Error | Error Type |  |  | System | Driver | Method |
| Pomeranz et al. 2021 | 0.12 | NA |  | -1.37 | -1.25 | Streams | Temperature | MLE |
| Yvon Durocher et al. 2011 (Community) | 0.13 | NA |  | -0.79 | -0.92 | FW mesocosms | Temperature | 10 ELB, not normalized |
| Dossena et al. 2012 (October) | 0.14 | 0.070 | se | NA | NA | FW mesocosms | Temperature | 6 ELB, not normalized |
| O’Gorman et al. 2017 | 0.15 | NA |  | -0.85 | -0.70 | Streams | Temperature | ln (mean species abundance) ~ ln(mean species mass) |
| Martinez et al. 2016 | 0.15 | NA |  | -1.11 | -0.96 | Streams | Land Use | 6 ELB, not normalized |
| Yvon Durocher et al. 2011 (Phytoplankton) | 0.19 | NA |  | -0.46 | -0.65 | FW mesocosms | Temperature | 10 ELB, not normalized |
| McGarvey and Kirk 2018 | 0.19 | NA |  | -1.81 | -1.62 | Streams | Seasonality | NAS, Log2 bins |
| Dossena et al. 2012 (April) | 0.21 | 0.064 | se | NA | NA | FW mesocosms | Temperature | 6 ELB, not normalized |

# kable() above looks awful  
# here is another bad option  
# if we decide to include this I'll make it look nice in excel.   
read.csv("lit comparison.csv")

## Author b\_diff error error\_type b\_low  
## 1 Pomeranz et al. 2021 0.12 NA -1.37  
## 2 Yvon Durocher et al. 2011 (Community) 0.13 NA -0.79  
## 3 Dossena et al. 2012 (October) 0.14 0.070 se NA  
## 4 O'Gorman et al. 2017 0.15 NA -0.85  
## 5 Martinez et al. 2016 0.15 NA -1.11  
## 6 Yvon Durocher et al. 2011 (Phytoplankton) 0.19 NA -0.46  
## 7 McGarvey and Kirk 2018 0.19 NA -1.81  
## 8 Dossena et al. 2012 (April) 0.21 0.064 se NA  
## b\_high System Driver  
## 1 -1.25 Streams Temperature  
## 2 -0.92 FW mesocosms Temperature  
## 3 NA FW mesocosms Temperature  
## 4 -0.70 Streams Temperature  
## 5 -0.96 Streams Land Use  
## 6 -0.65 FW mesocosms Temperature  
## 7 -1.62 Streams Seasonality  
## 8 NA FW mesocosms Temperature  
## Method  
## 1 MLE  
## 2 10 ELB, not normalized  
## 3 6 ELB, not normalized  
## 4 ln (mean species abundance) ~ ln(mean species mass)  
## 5 6 ELB, not normalized  
## 6 10 ELB, not normalized  
## 7 NAS, Log2 bins  
## 8 6 ELB, not normalized