Supporting Information. Foster, S.D., J. Vanhatalo, V.M. Trenkel, T. Schulz, E. Lawrence, R. Przeslawski, and G.R. Hosack. 2021. Effects of ignoring survey design information for data reuse. Ecological Applications.

Appendix S1

Section S1: Covariate Coverage

To aid interpretation of the simulation experiment in Methods: Simulation Experiment, we investigate how the simulated surveys sample the covariates in response to varying α. To this end, we simulated designs according to the process described in Methods: Simulation Experiment and calculate the sampled range of the covariates. This is then compared to the range of the covariates in all the study region, using the statistic $100^{\text{range(sample)}}/_{\text{range(studyarea)}}$. We choose to include larger sample sizes, much larger than realistic, to see how much sampling is required for good covariate coverage.

The results are in Figure S1. Increasing α decreases coverage of bathymetry – very deep habitats are biased against. However, increasing α increased coverage of TPI. The inclusion prob-abilities also bias away from unfavorable TPI values, but the favorable ones are quite rare and patchy. The combined effect is that the small α values do not sample these TPI habitats, whereas the larger α values do but still capture enough of the unfavorable habitat.

It is important to note that there are N = 8840 potential survey locations with positive inclusion probabilities from the original survey design. When α is increased, some of these will be assigned to zero (see Methods: Simulation Experiment). Some of these sample sizes and some of the α values are quite extreme, and generate situations where more samples are required to be drawn than there are locations (with non-zero inclusion probabilities). Obviously, this cannot be done, and so these combinations are not plotted.

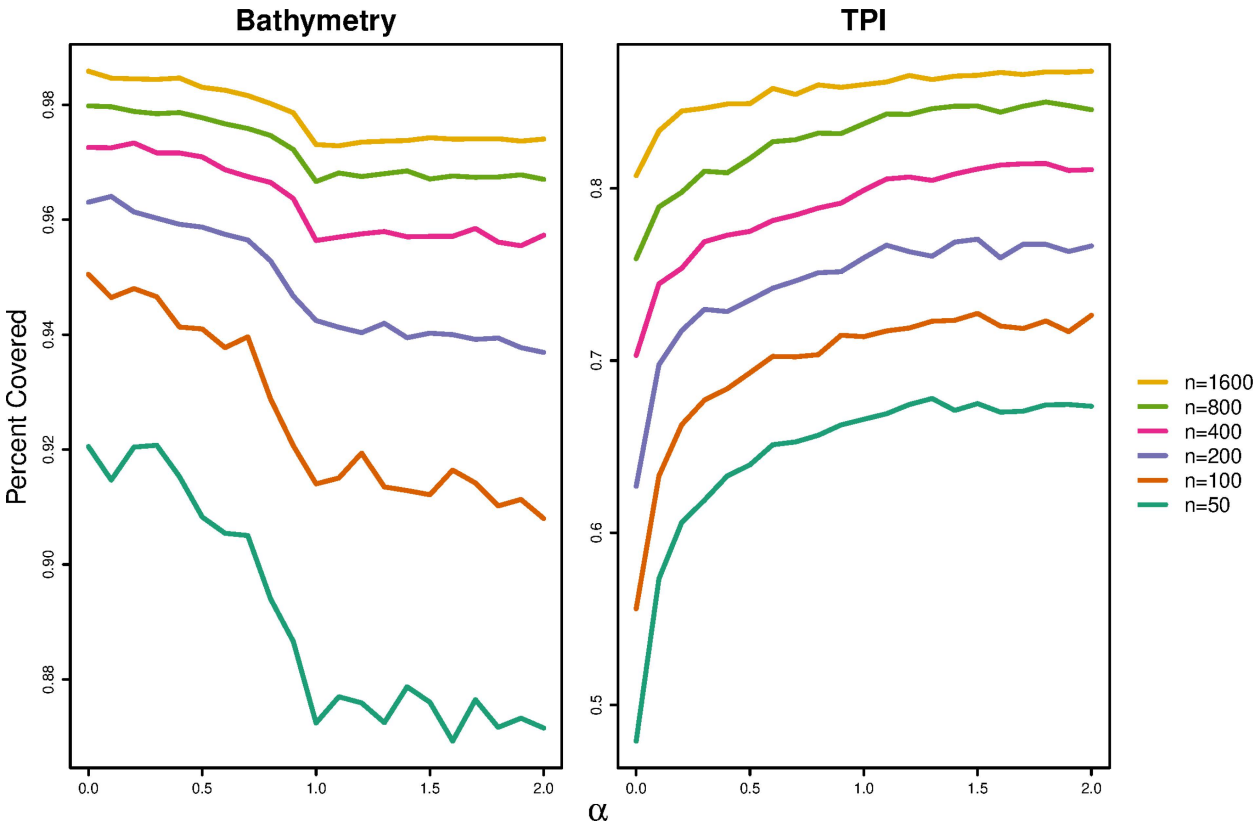

Figure S1. Proportion of covariate space covered with surveys. Average of $B = 250$ surveys. The statistic is each covariate's sampled range divided by its actual range in the survey area. Low values give poorer coverage. Small values of α give more even inclusion probabilities.

Section S2: Extra Plots for Seamount Analysis

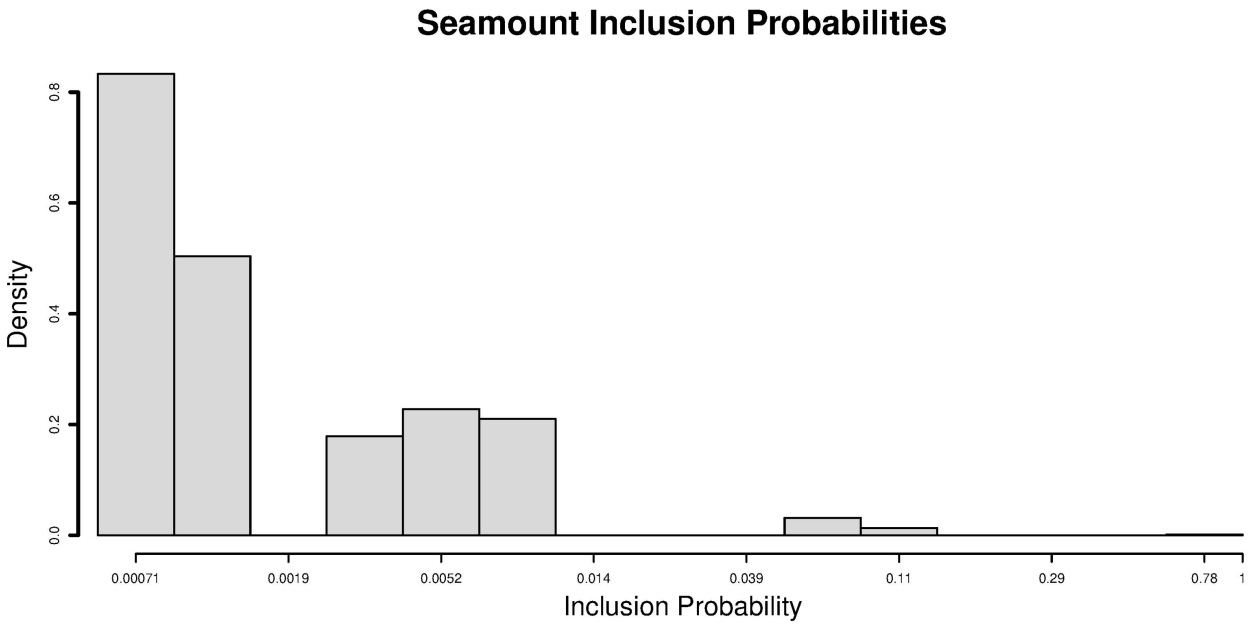

Figure S2: Distribution of the inclusion probabilities used to create designs for the seamount sampling area. There are 14 distinct values that correspond to the bathymetry and TPI categories in Foster et al. (2020).

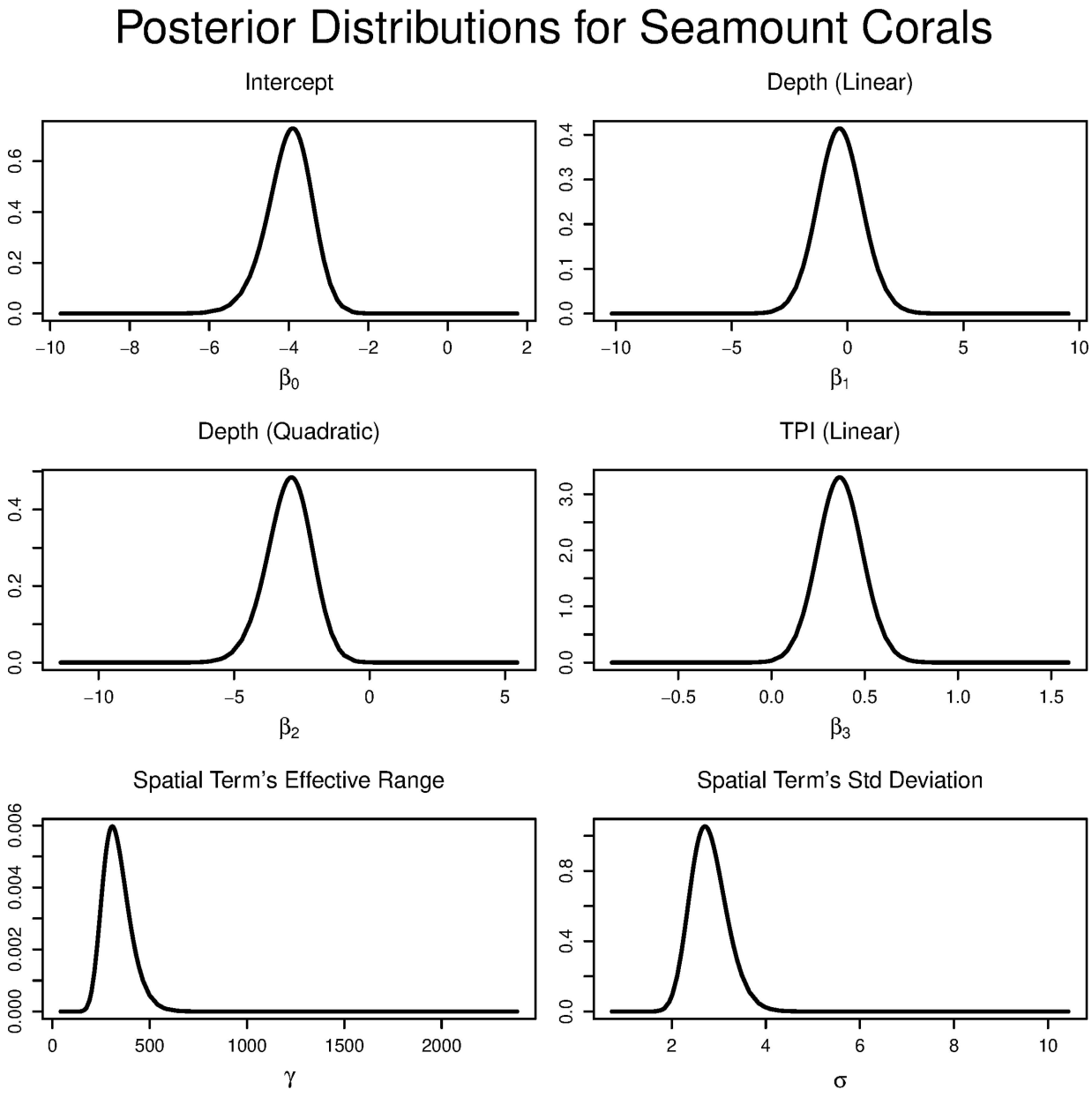

Figure S3: Posterior distributions of the parameters from the spatial model for the corals. See Eq. 1 in Methods: A Model for Coral Distribution, and subsequent text, for parameter definitions.

References

Foster, S. D., G. R. Hosack, J. Monk, E. Lawrence, N. S. Barrett, A. Williams, and R. Przeslawski. 2020. Spatially balanced designs for transect-based surveys. Methods in Ecology and Evolution 11:95–105.