SUPPLEMENTARY INFORMATION

Effects of sample size

The simulations described in the main text were repeated for sample sizes n = 30, 100 and 1000 to test for potential differences owing to sample size. The same pattern persists throughout the different sample sizes (Tables S1 and S2) – SRM Type I error rates are all ~5% and the estimated slope β is never significantly different from the input parameter b = 0.6, while SDDM Type I error rates are greater than 5%, increasing with residual error (σ_e) and β is consistently significantly different from b = 0.6. SDDM Type I error rates tend to be lower at smaller n (as low as ~12%), and the detrimental effects of decoupling a paired, bivariate dataset are more pronounced and readily detected in larger n (approaching 100% for the slope estimate). However, this *does not* mean that statistical consequences are less severe for smaller datasets, but simply that they are less readily detected in smaller datasets – the problem still persists. Therefore, across all sample sizes and residual error, the SDDM is demonstrated here as not being statistically viable as a base model to correct for sampling bias.

SUPPLEMENTARY FIGURES

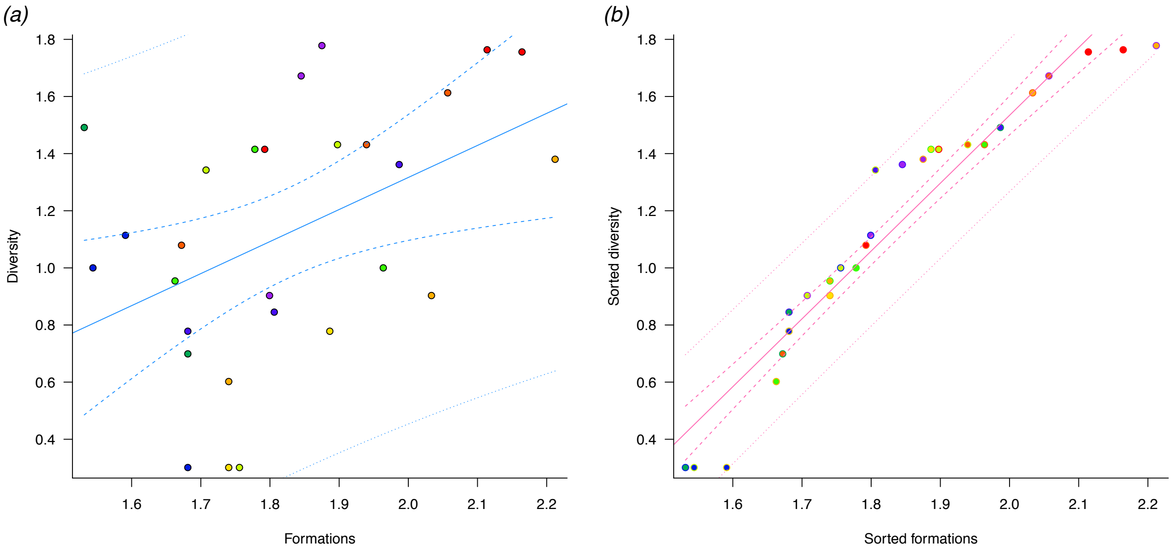

Figure S1. The difference between the original paired, bivariate relationship (a) and the forced, false relationship (b) shown using the data from Benson et al. (2010). Log-transformed marine tetrapod diversity has a significant but weak relationship with log-transformed formation count ($r^2 = 0.206$, $p = 0.0152$; a). However, once diversity and formation count are sorted independently of each other following Smith and McGowan (2007), then the relationship becomes significant and strong ($r^2 = 0.924$, $p < 0.001$; b). Points are coloured according to their geological age with cooler colours on the older and warmer colours on the younger ends of the time scale. Filled and outline colours in (b) correspond to the ages of the rock record and diversity respectively, and demonstrate visually the mismatch between y' and x'. Dashed lines are confidence intervals while dotted lines are prediction intervals.

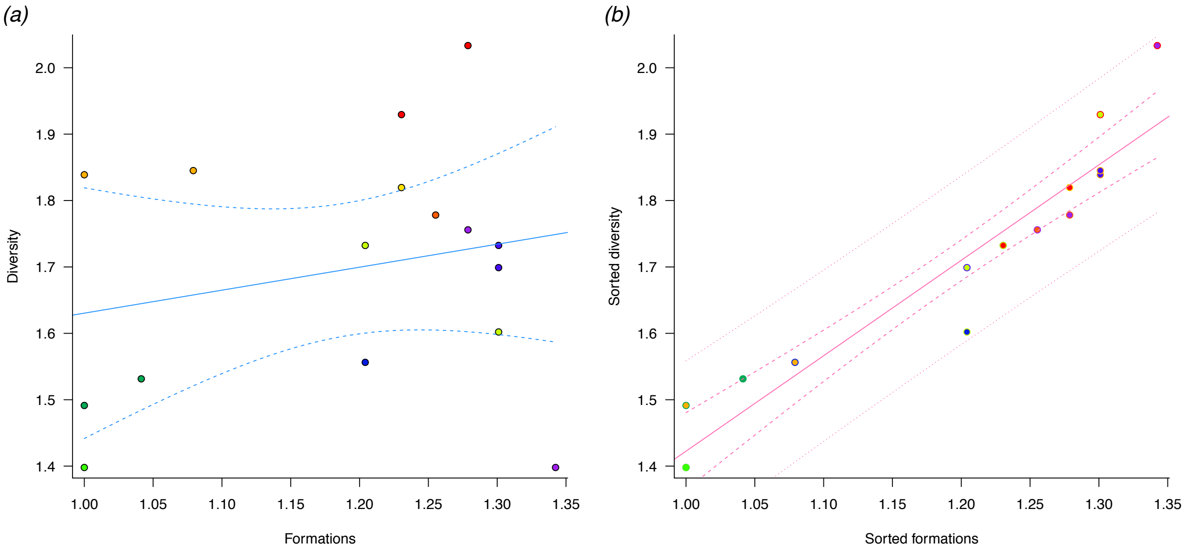

Figure S2. The difference between the original paired, bivariate relationship (a) and the forced, false relationship (b) shown using the data from Benson and Upchurch (2013). Log-transformed early Permian tetrapod diversity has a non-significant and weak relationship with log-transformed formation count (r^2 = 0.05, p = 0.392; a). However, once diversity and formation count are sorted independently of each other following Smith and McGowan (2007), then the relationship becomes significant and strong (r^2 = 0.9, p < 0.001; b). Points are coloured according to their geological age with cooler colours on the older and warmer colours on the younger ends of the time scale. Filled and outline colours in (b) correspond to the ages of the rock record and diversity respectively, and demonstrate visually the mismatch between y' and x'. Dashed lines are confidence intervals while dotted lines are prediction intervals.

SUPPLEMENTARY TABLES

Table S1. Type I error rates (%) for SRM and SDDM estimates (intercept α and slope β) across sample size (n) and residual error (σ_e).

N	σ_e	**SRM**		**SDDM**	
		α	β	α	β
30	0.05	4.88	5.44	12.1	12.4
	0.10	5.18	4.88	20.5	19.4
	0.25	5.36	5.32	36.8	50.4
	0.50	5.46	5.60	52.0	87.3
100	0.05	5.34	4.90	26.1	28.5
	0.10	4.84	4.92	40.2	48.4
	0.25	4.82	4.78	57.3	91.3
	0.50	5.48	5.14	68.7	100.0
1000	0.05	4.82	5.26	65.7	83.6
	0.10	4.54	4.64	74.6	99.0
	0.25	4.80	4.80	83.6	100.0
	0.50	5.06	4.88	88.9	100.0

Table S2. t-test results between mean regression slopes of 5000 iterations and the theoretical slope b = 0.6, for SRM and SDDM across sample size (n) and residual error (σ_e).

n	σ_e	SRM			SDDM		
		mean-slope	t-value	p-value	mean-slope	t-value	p-value
30	0.05	0.6	0.494	0.621	0.601	6.6	0
	0.10	0.6	-0.524	0.600	0.605	18.2	0
	0.25	0.6	-1.080	0.282	0.640	59.4	0
	0.50	0.6	1.370	0.171	0.768	123.0	0
100	0.05	0.6	1.230	0.220	0.602	20.9	0
	0.10	0.6	-1.790	0.073	0.607	46.0	0
	0.25	0.6	-0.042	0.967	0.646	131.0	0
	0.50	0.6	0.685	0.493	0.775	244.0	0
1000	0.05	0.6	0.037	0.971	0.602	88.2	0
	0.10	0.6	-0.563	0.573	0.608	181.0	0
	0.25	0.6	0.689	0.491	0.650	446.0	0
	0.50	0.6	0.601	0.548	0.780	805.0	0

LITERATURE CITED IN THE SI:

Benson, R.B.J., Butler, R.J., Lindgren, J. & Smith, A.S. (2010) Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates. *Proceedings Of The Royal Society B-Biological Sciences,* **277,** 829-834.

Benson, R.B.J. & Upchurch, P. (2013) Diversity trends in the establishment of terrestrial vertebrate ecosystems: Interactions between spatial and temporal sampling biases. *Geology,* **41,** 43-46.

Smith, A.B. & McGowan, A.J. (2007) The shape of the phanerozoic marine palaeodiversity curve: How much can be predicted from the sedimentary rock record of western Europe? *Palaeontology,* **50,** 765-774.