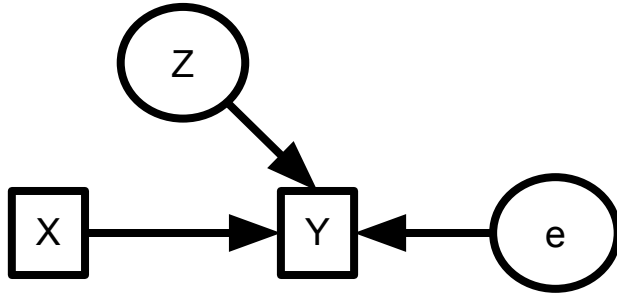**A. Uncorrelated Unmeasured Driver Causing Only Additional Error**

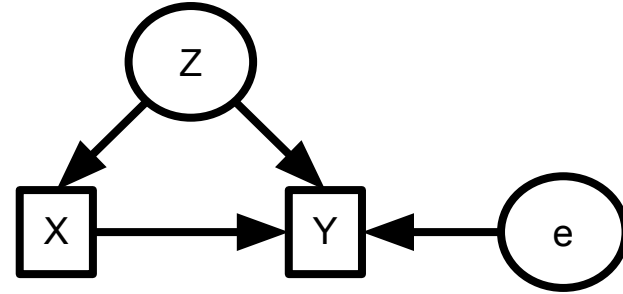**B. Correlated Driver Causing Omitted Variable Bias**

Figure 1: A response variable of interest (Y) is driven by both a measured variable (X) and an unmeasured variable (Z) as well as other uncorrelated sources of residual variability (e). In one scenario, X and Z are uncorrelated, and thus the lack of inclusion of Z in a statistical model increases the standard error of the estimate of the effect of X on Y, but does not alter the causal identification of the statistical model. If Z drives X, however (or if Z and X are driven by a common unmeasured driver), the omission of Z from a statistical model causes bias in the estimate of X on Y in a direction that depends on the effect of Z on Y. The results of the model are no longer causally identified.
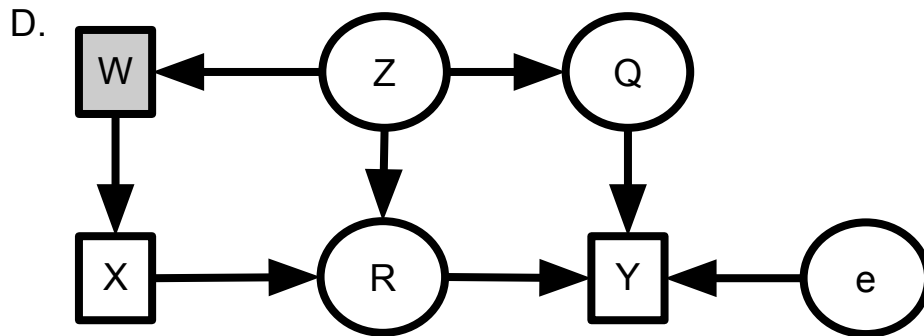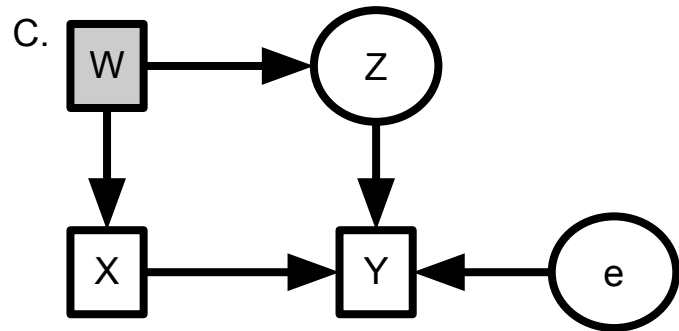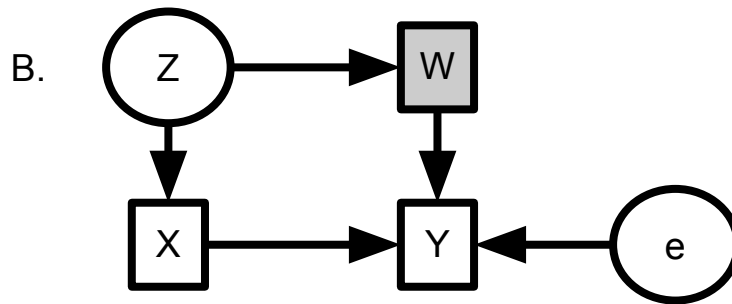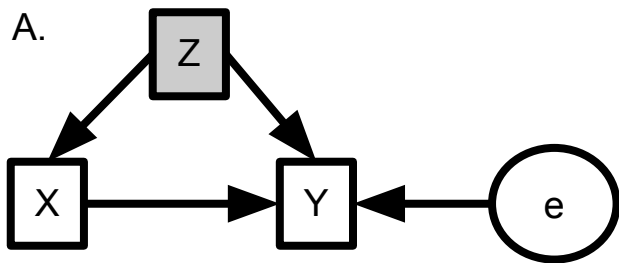
Figure 2: Examples of satisfying the back-door criterion with control variables. By including shaded observed variables, either Z or W, in a statistical analysis of the effects of X on Y, omitted variable bias is controlled for the results have a causal interpretation. The relationship between the control variable and Y might (A and B) or might (C and D) not have a causal interpretation, depending on the structure of the system. Note, in (D), Q would have also served as an adequate control instead of W. R would have been a bad control.
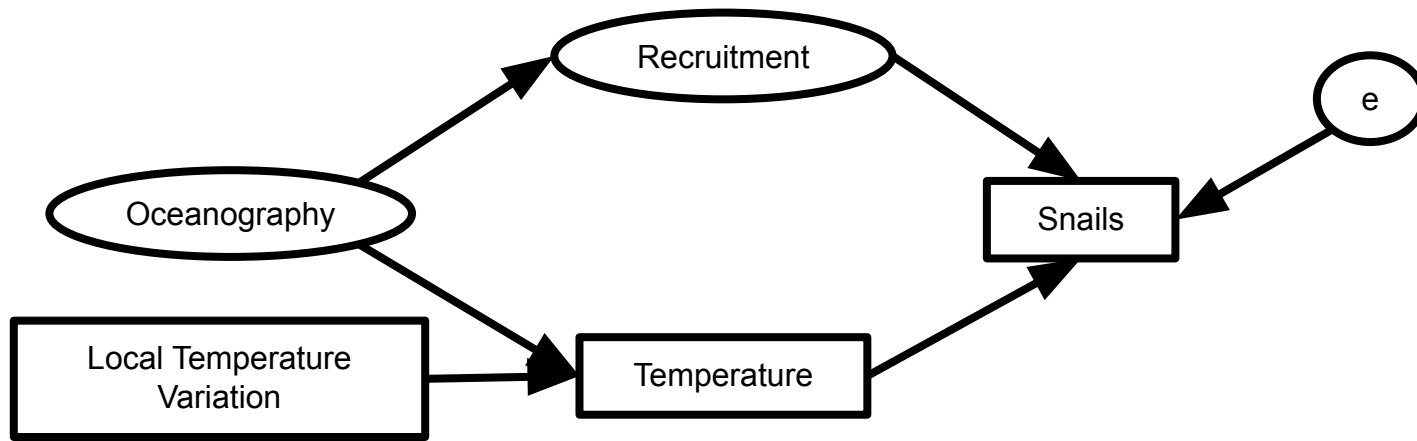
Figure 3: A system describing the controls of snail abundance in the intertidal. Oceanography regulates both temperature and recruitment, both of which drive snail abundance. Temperature, however, is also driven by local influences as well. This could be variability in plot-level temperature within a site or site-level temperature variability between years uncorrelated with local oceanography and recruitment.

A. Path Diagram of a Mixed Model with Site Random Intercept
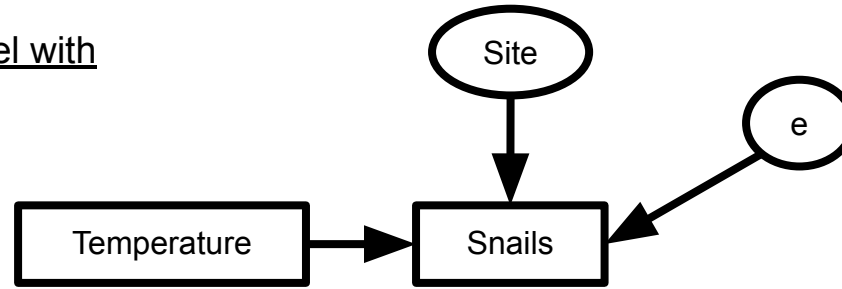
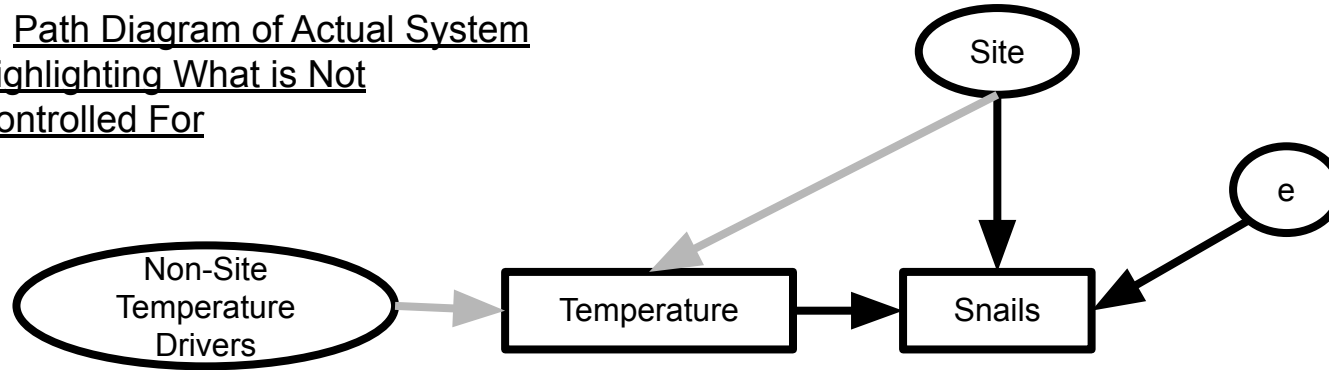B. Path Diagram of Actual System Highlighting What is Not Controlled For

Figure 4: The system assumed underlying a mixed model (A) versus the system as it is (B). Note that a mixed model does not account for the correlation between site and temperature nor does it separate out the non-site drivers of temperature. Instead, the effect of temperature on snails is confounded by any correlated site-level drivers that correlate with temperature at the site level.
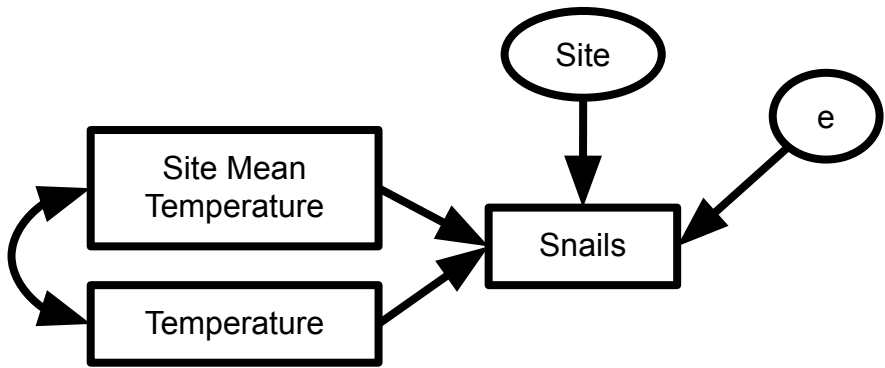
A. Fixed Effects Transformation Path Model
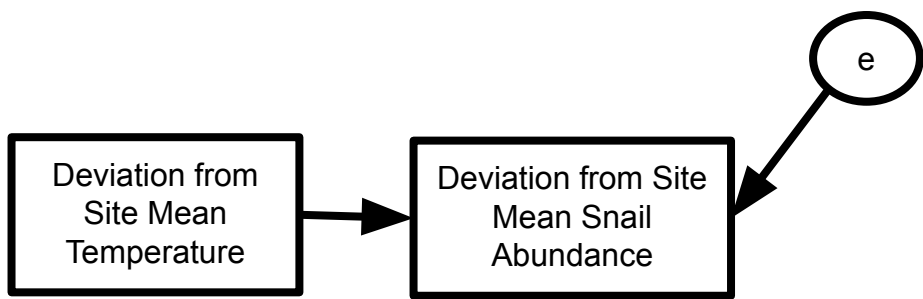
B. Fixed Effects with a Site Dummy Variable Path Model
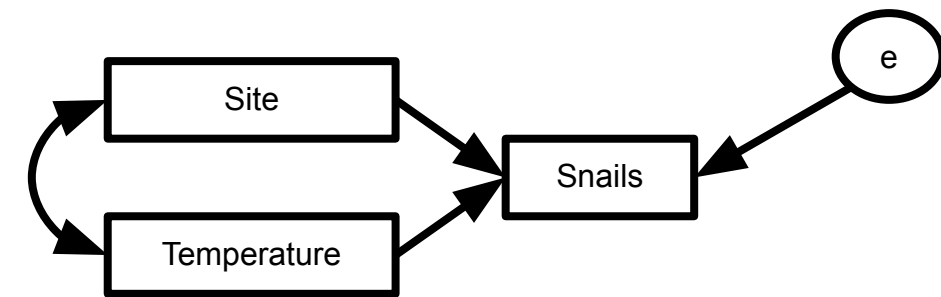
C. Group Mean Covariate Path Model
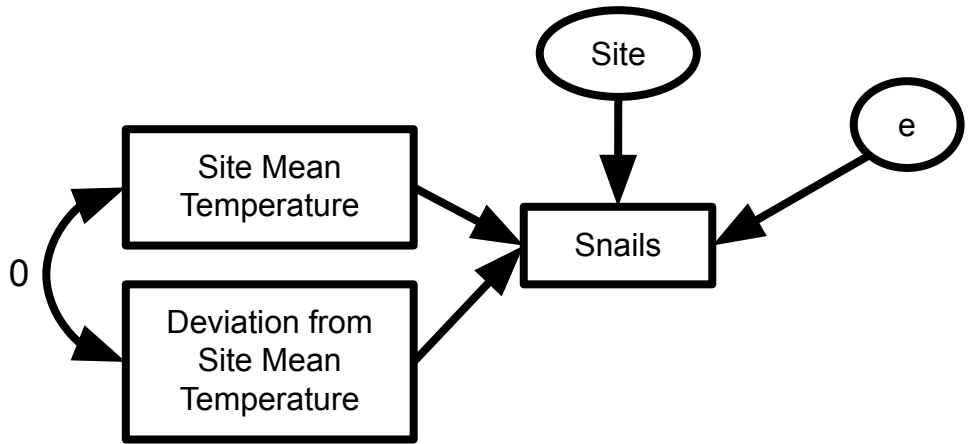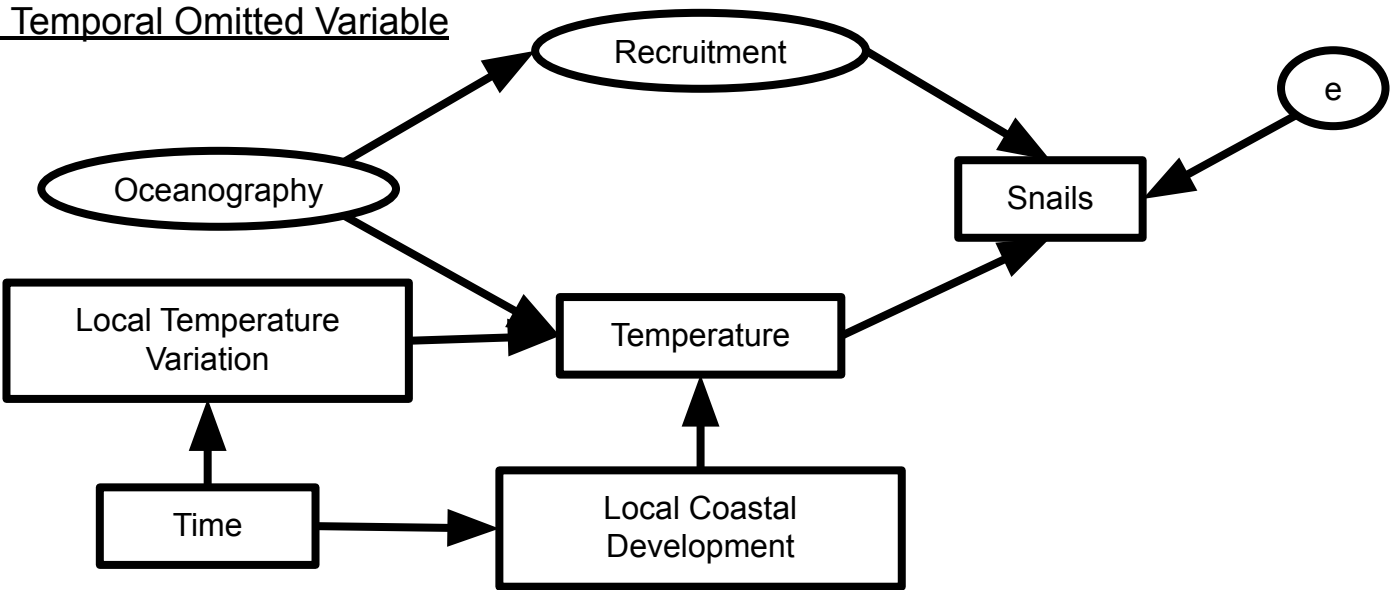
D. Group Mean Centered Path Model
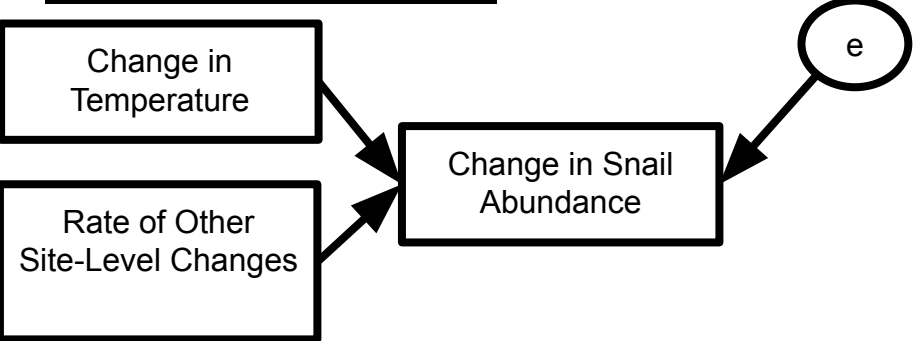
Figure 5: Path model representations of different statistical models handling spatial omitted variables in the text.
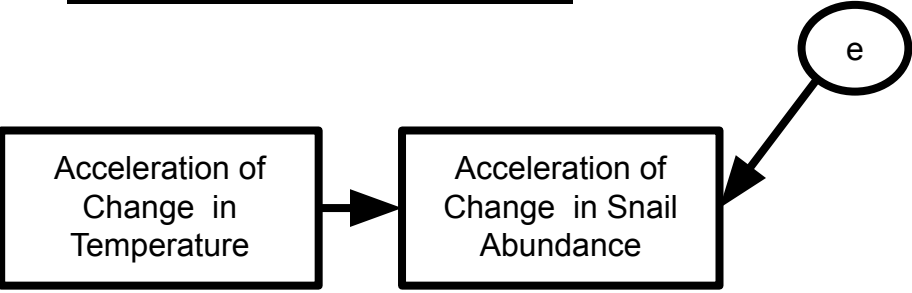
Figure 6: Path model representations of different models handling spatial and temporal omitted variables.
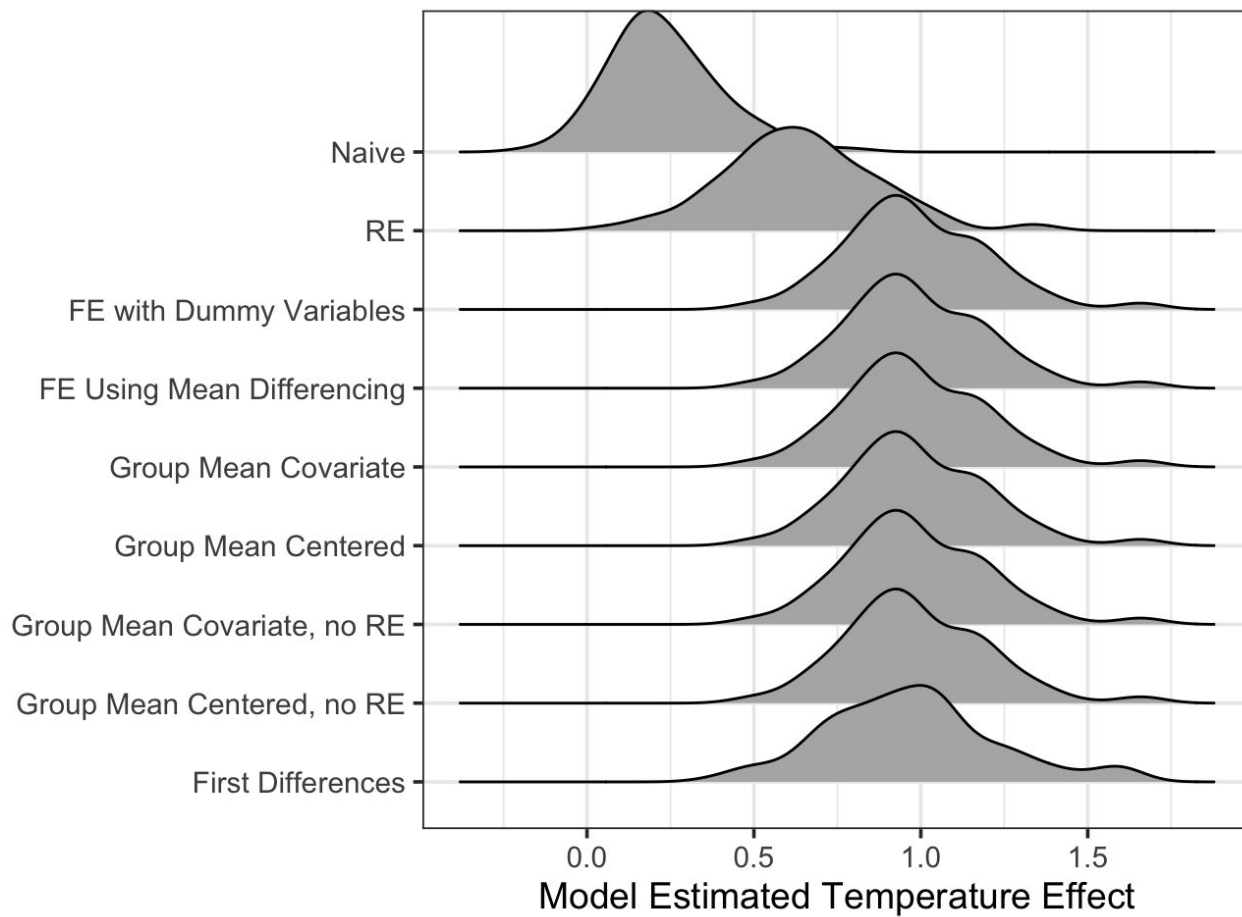
Figure 7: Distribution of point estimates of temperature effects from different models.
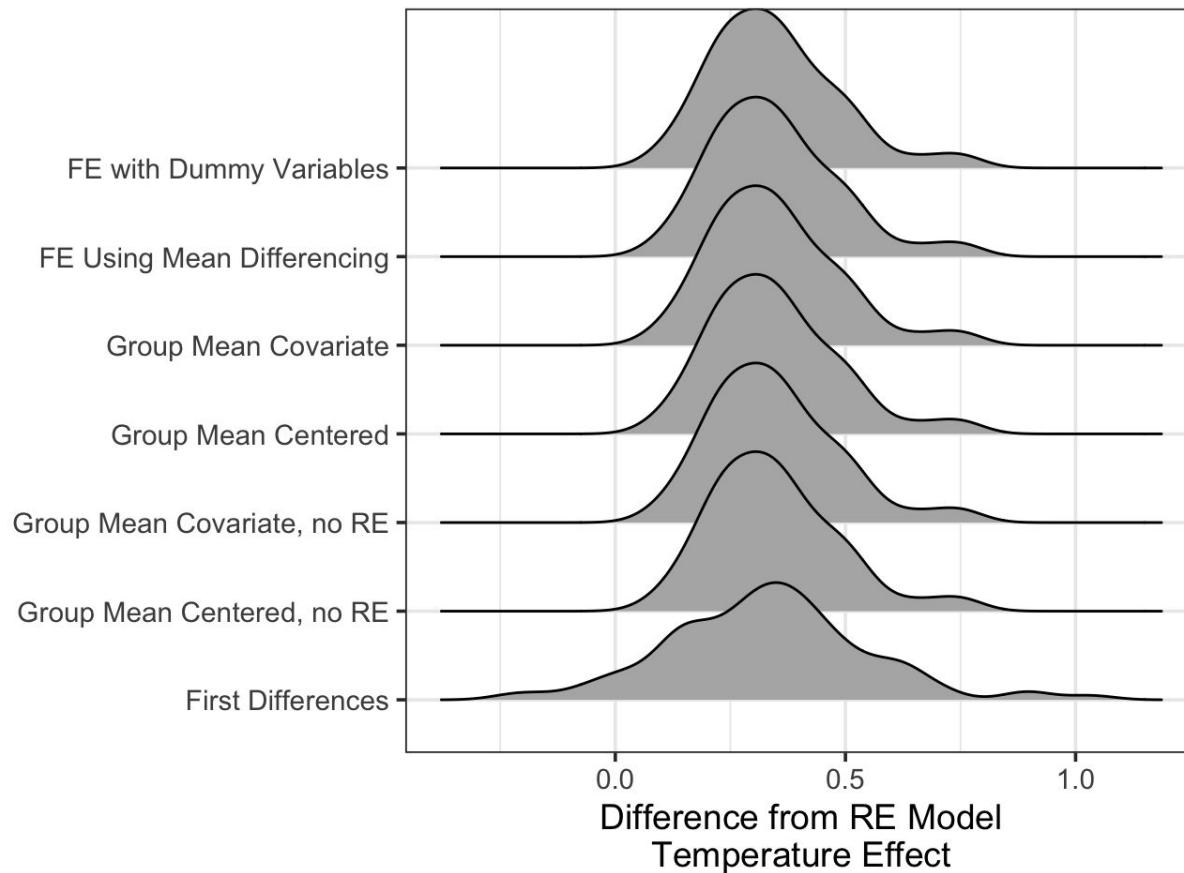
Figure 8: Distribution of difference between point estimate of temperature effect in the RE model versus all other models demonstrating clear bias in the estimand of the RE model.

Table 1: Mean and SD of point estimates of temperature effects from different models.

Model Type	Mean Estimate	SD Estimate
Naive	0.231	0.165
RE	0.640	0.232
FE with Dummy Variables	0.985	0.215
FE Using Mean Differencing	0.985	0.215
Group Mean Covariate	0.985	0.215
Group Mean Centered	0.985	0.215
Group Mean Covariate, no RE	0.985	0.215
Group Mean Centered, no RE	0.985	0.215

Table 2: Fraction of simulated runs where the mean +/- 2 SE of the temperature effect either overlapped 0 (i.e., high likelihood of committing a type II error) or did *not* contain the true effect of temperature.

Model Type	95% CI Contains 0	95% CI does Not Contain 1
RE	0.08	0.54
FE with Dummy Variables	0.00	0.05
FE Using Mean Differencing	0.00	0.05
Group Mean Covariate	0.00	0.05
Group Mean Centered	0.00	0.05
Group Mean Covariate, no RE	0.01	0.04
Group Mean Centered, no	0.01	0.04