Decompositions

Sophie Woodward

2023-06-08

Contents

Plot eigenvectors by eigenvalue

Key Findings (CHANGES NEEDED)

- When the data-generating mechanism is based on the **nested** decomposition, the outcome model is **linear**, and confounding dissipates **locally** (within 5×5 grids), both the nested and spectral decompositions recover unbiased estimates of the treatment effect at small spatial scales.
- When the data-generating mechanism is based on the **spectral** decomposition, the outcome model is **linear**, and confounding dissipates **locally** (at high spectral frequencies), the spectral decomposition recovers unbiased estimates of the treatment effect at small spatial scales. If the spectral correlation of exposure and confounder is equal to zero up to a small-enough frequency (big-enough spatial scale) the nested decomposition recovers nearly unbiased estimates at small spatial scales as well.
- When the data-generating mechanism is based on the **nested** decomposition, the outcome model is **linear**, and confounding dissipates **globally** (confounding within 5×5 grids but not across), I thought the plots would be the same as the previous, but x axis (spatial scale) flipped. They are not exactly.
- When there is an **interaction** between X and Z but confounding still dissipates locally, there is still near-zero bias at small spatial scales under both DGMs. At higher spatial scales bias is worse.
- Neither decomposition can recover unbiased estimates at any scale when there is a **quadratic** term of exposure X is included in the outcome model.

Note: in the following plots, I mark the x axis by spatial scale. If I plot results from the nested decomposition, then there are only two points: spatial scale equal to 1 is the so-called county level (1×1 grid) and spatial scale equal to 2 is the so-called state level (5×5 grid). If I plot results from the spectral decomposition, then the spatial scale indexes the ordered eigenvalues of the graph Laplacian. So the spatial scales between spectral and nested plots should not be directly compared.

Simulation 1: Nested DGM

Denote X as exposure X, Z as unmeasured confounder, and Y as outcome. We simulate the following scenario 100 times.

1. Generate $n = 729$ observations of $Z \sim \text{Expo}(1)$.
2. Decompose Z using the nested decomposition with three levels:

$$Z(1) = E(Z|L_1),$$
$$Z(2) = E(Z|L_2) - Z(1)$$
$$Z(3) = Z - Z(1) - Z(2)$$

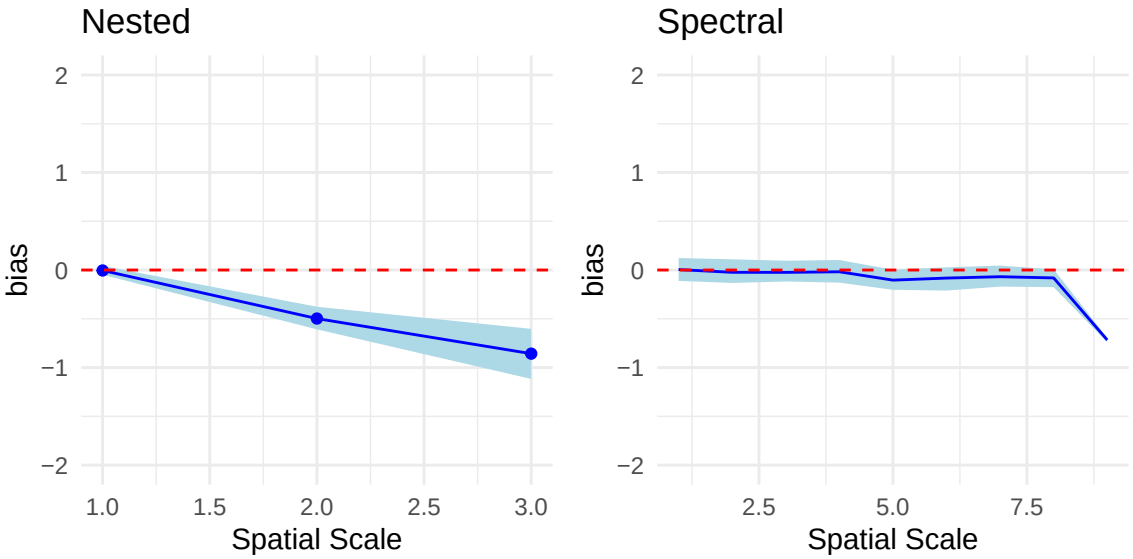
   where L_1 is a random variable that maps each of the 729 units to their corresponding 9×9 grid (call this a state) and L_2 is a random variable that maps each of the units to their corresponding 3×3 grid (call this a county). Note that $Z = Z(1) + Z(2) + Z(3)$.
3. Repeat the above to generate independent noise: generate $n = 729$ observations of $\zeta \sim \text{Expo}(1)$. Decompose ζ using the nested decomposition with three levels to obtain $\zeta = \zeta(1) + \zeta(2) + \zeta(3)$.
4. At each level $l = 1, 2, 3$ create $X(i)$ with the following formula:

$$X(i) = \rho_i Z(i) + \sqrt{1 - \rho_i^2}\zeta(i)$$

   using $\rho = (0.9, 0.5, 0.001)$.
5. Construct X using the formula $X = X(1) + X(2) + X(3)$.
6. Let $Y = 2X - Z + \epsilon$ where $\epsilon \sim \mathcal{N}(0, 1)$ independently.

By construction, X, Z are nearly uncorrelated within the counties of 3×3, but correlated both within the states of 9×9 and across states.

For each of the 100 scenarios, we decompose X, Z, Y at different spatial scales using 1) nested decomposition and 2) spectral decomposition. At each spatial scale ω, we obtain an estimate $\hat{\beta}(\omega)$ of $\beta = 2$ from a linear regression of $Y(\omega)$ on $X(\omega)$. Hypothesis: $\hat{\beta}(\omega)$ is unbiased at ω corresponding to finer spatial scales since by construction confounding dissipates locally.



Simulation 2: Spectral DGM

We repeat the following but using the spectral decomposition to generate data.

1. Generate $n = 729$ observations of $Z \sim \text{Expo}(1)$.
2. Project Z into the spectral domain using the spectral decomposition:

$$Z^* = G^T Z$$

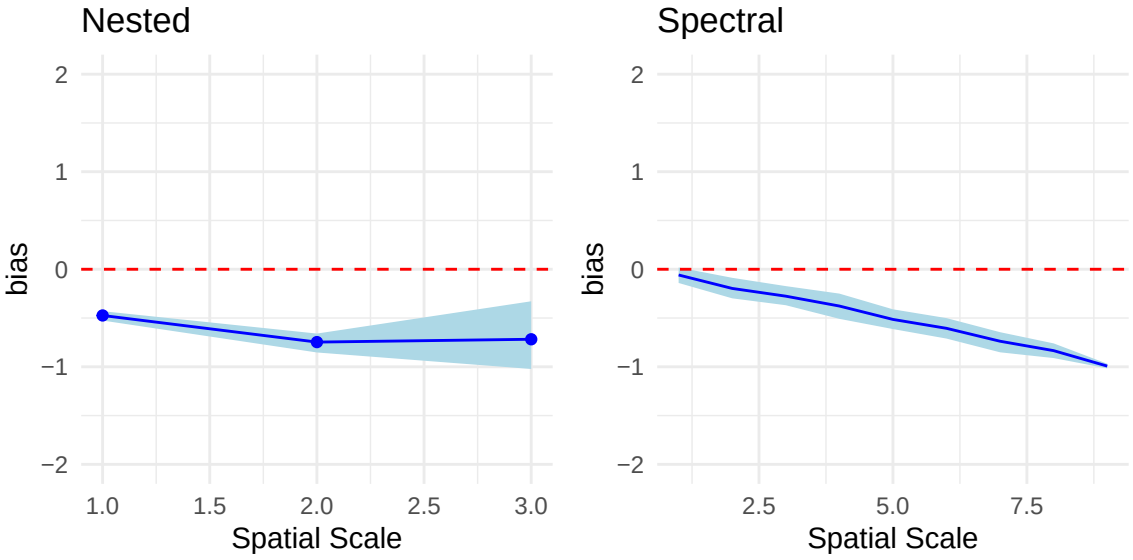
   where G is the graph Laplacian.

3. Repeat the above to generate independent noise: generate $n = 729$ observations of $\zeta \sim \text{Expo}(1)$. Project ζ into the spectral domain to obtain $\zeta^* = G^T \zeta$.
4. At each spectral frequency ω create X^* with the following formula:

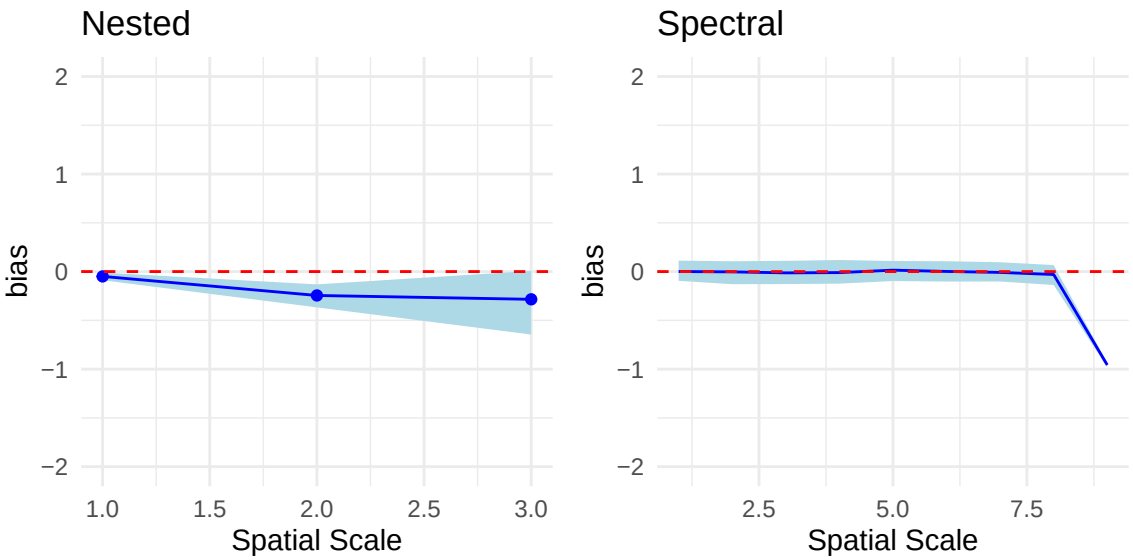
$$X^*(\omega) = \rho_i Z^*(\omega) + \sqrt{1 - \rho_i^2}\zeta^*(\omega)$$

   using $\rho = (1, 727/728, \ldots, 1/728, 0)$. The frequencies ($n = 729$) are sorted lowest to highest, so that spatial scale is sorted highest to lowest. By construction, X and Z become less correlated at smaller scales.
5. Construct X using the formula $X = GX^*$.
6. Let $Y = 2X - Z + \epsilon$ where $\epsilon \sim \mathcal{N}(0, 1)$ independently.

Again, for each of the 100 scenarios we decompose X, Z, Y at different spatial scales using 1) nested decomposition and 2) spectral decomposition. At each spatial scale ω, we obtain an estimate $\hat{\beta}(\omega)$ of $\beta = 2$ from a linear regression of $Y(\omega)$ on $X(\omega)$. Hypothesis: $\hat{\beta}(\omega)$ is unbiased at ω corresponding to finer spatial scales since by construction confounding dissipates locally.



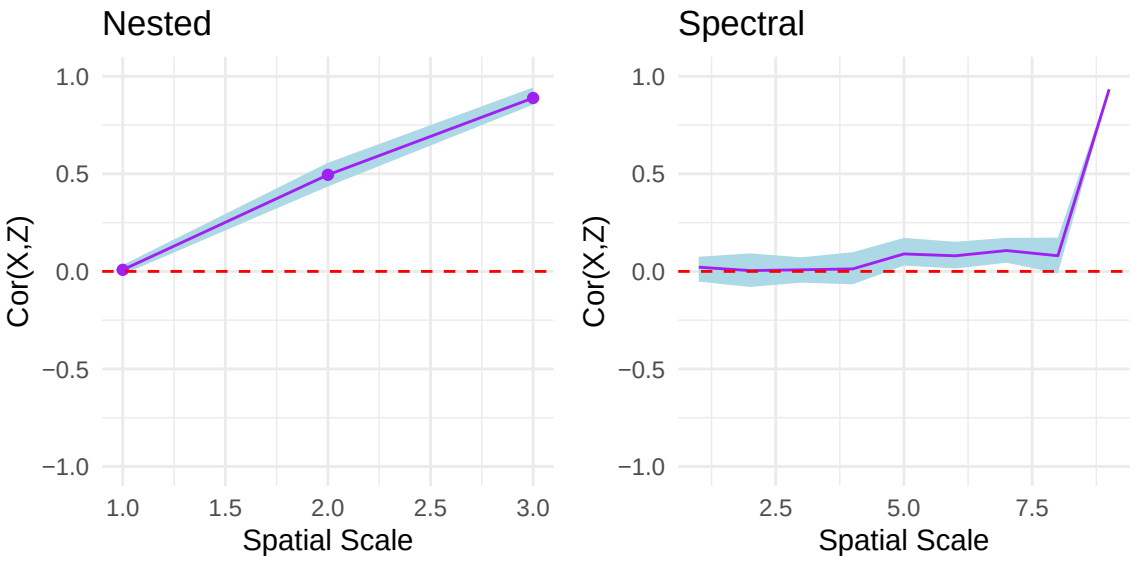
Looking at the spectral plot, it's reassuring to see that the bias of the coefficient is 0 at low spatial scales when the DGM is indeed spectral. The estimates obtained from the nested decomposition are biased at both of the two grid levels. I think this makes sense: by construction the covariance is continuously decreasing with spatial scale; within a 3×3 grid we will observe bias. Let's try $\rho_i = 0$ for all but the 100 lowest frequencies (100 largest scales).



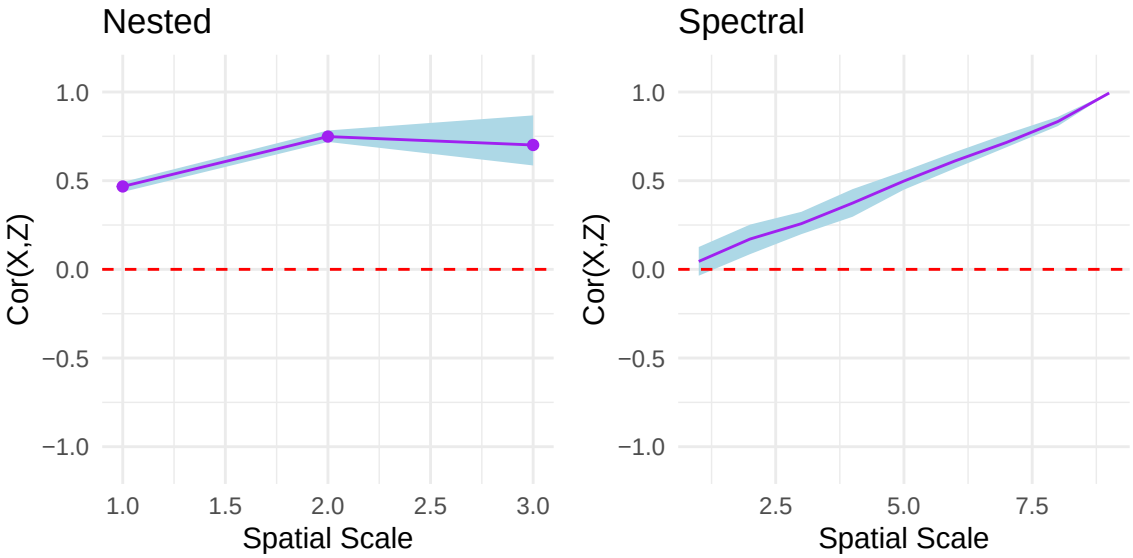
We observe less bias.

Simulation 3: Correlation by Spatial Scale

Let's plot the correlation between X and Z by spatial scale for both decompositions using the setup from simulation 1 (nested decomposition, $\rho = (0.9, 0.5, 0.001)$).

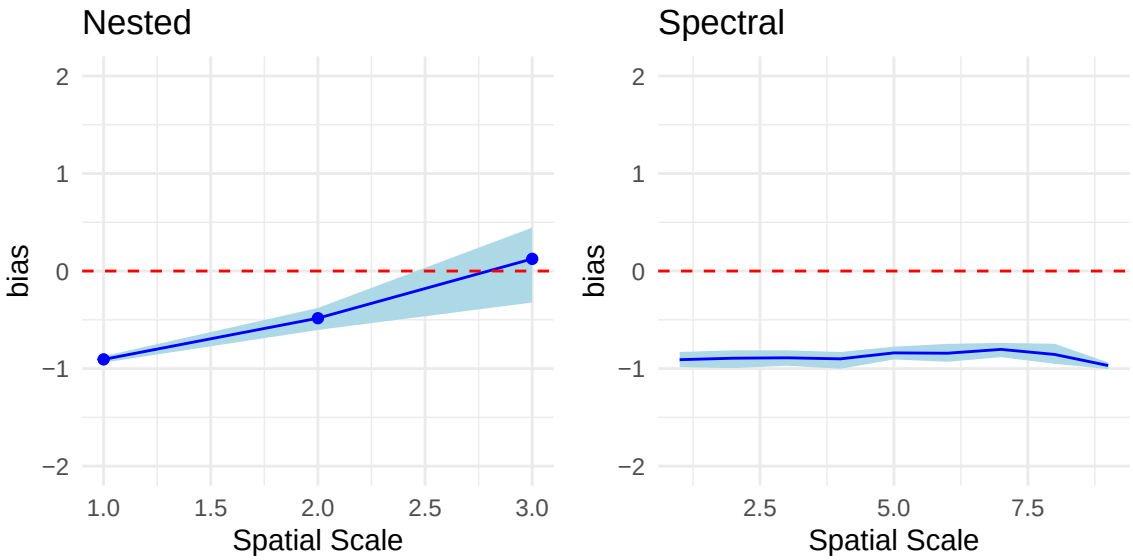


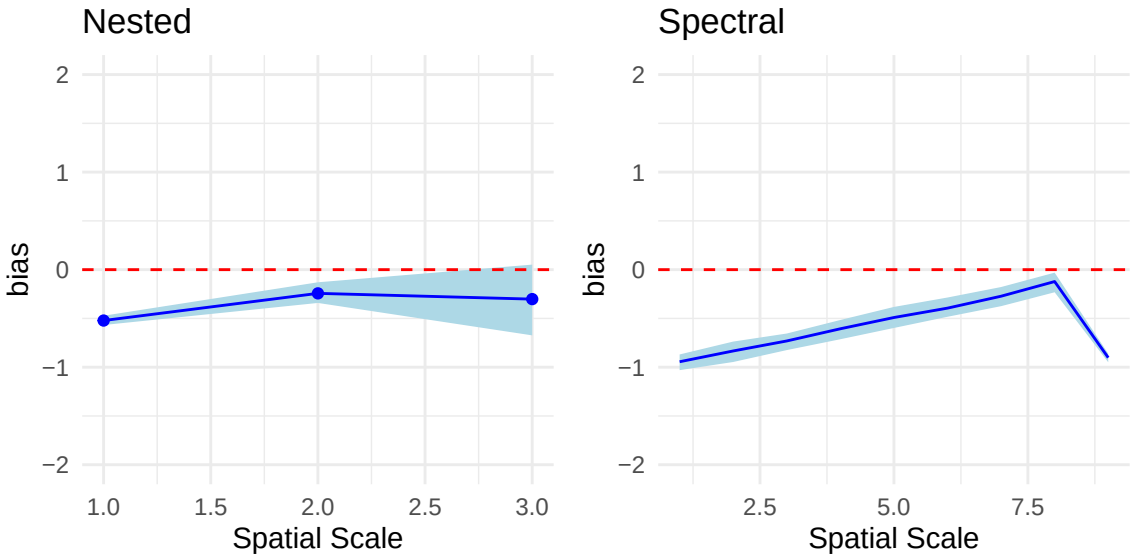
Now, let's plot the correlation between X and Z by spatial scale for both decompositions using the setup from simulation 2 (spectral decomposition, $\rho = (1, 727/728, \ldots, 1/728, 0)$).



Simulation 4: Local Confounding

We repeat simulations 1 and 2 but now confounding dissipates at larger scales rather than smaller ones.

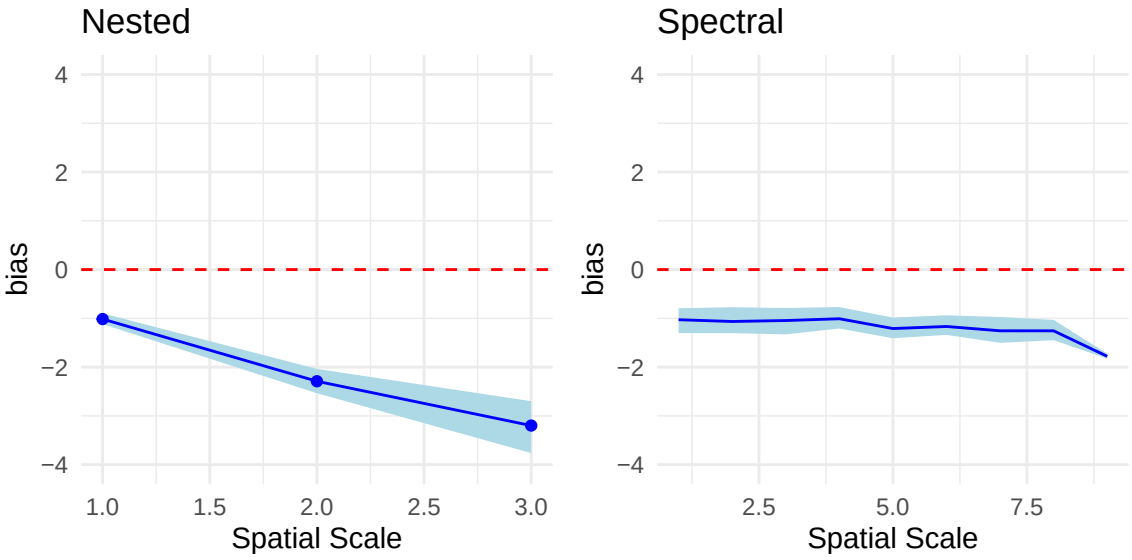




Simulation 5: Outcome Model with Interaction

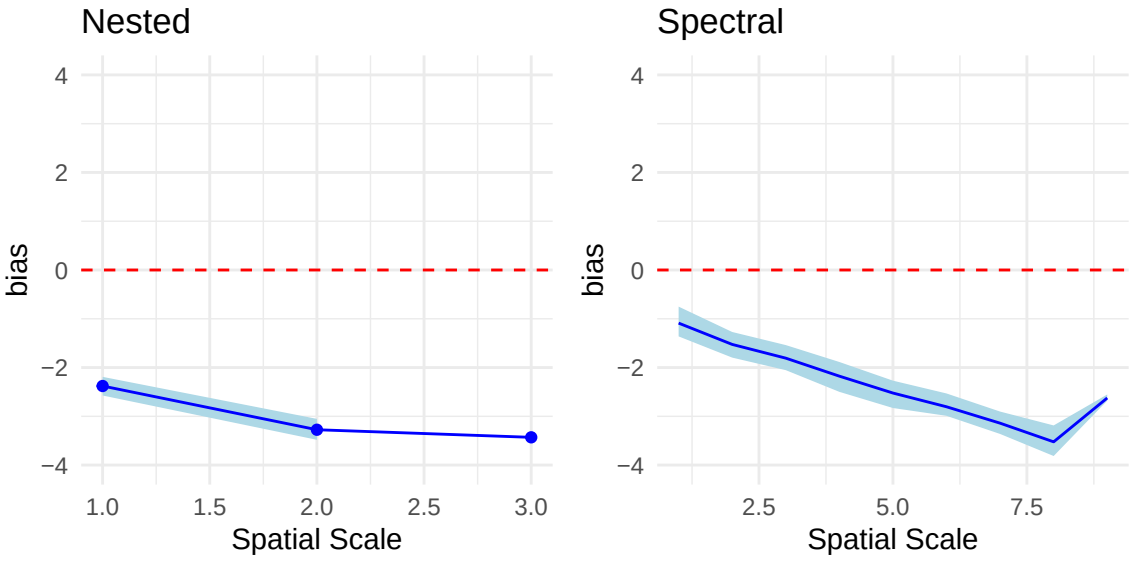
We repeat simulation 1 and 2 but the outcome model now includes an interaction term between X and Z. In particular, let $Y = 2X - Z - XZ + \epsilon$.

Nested DGM



The bias of the estimates is worse at higher scales but still zero bias at low scales.

Spectral DGM

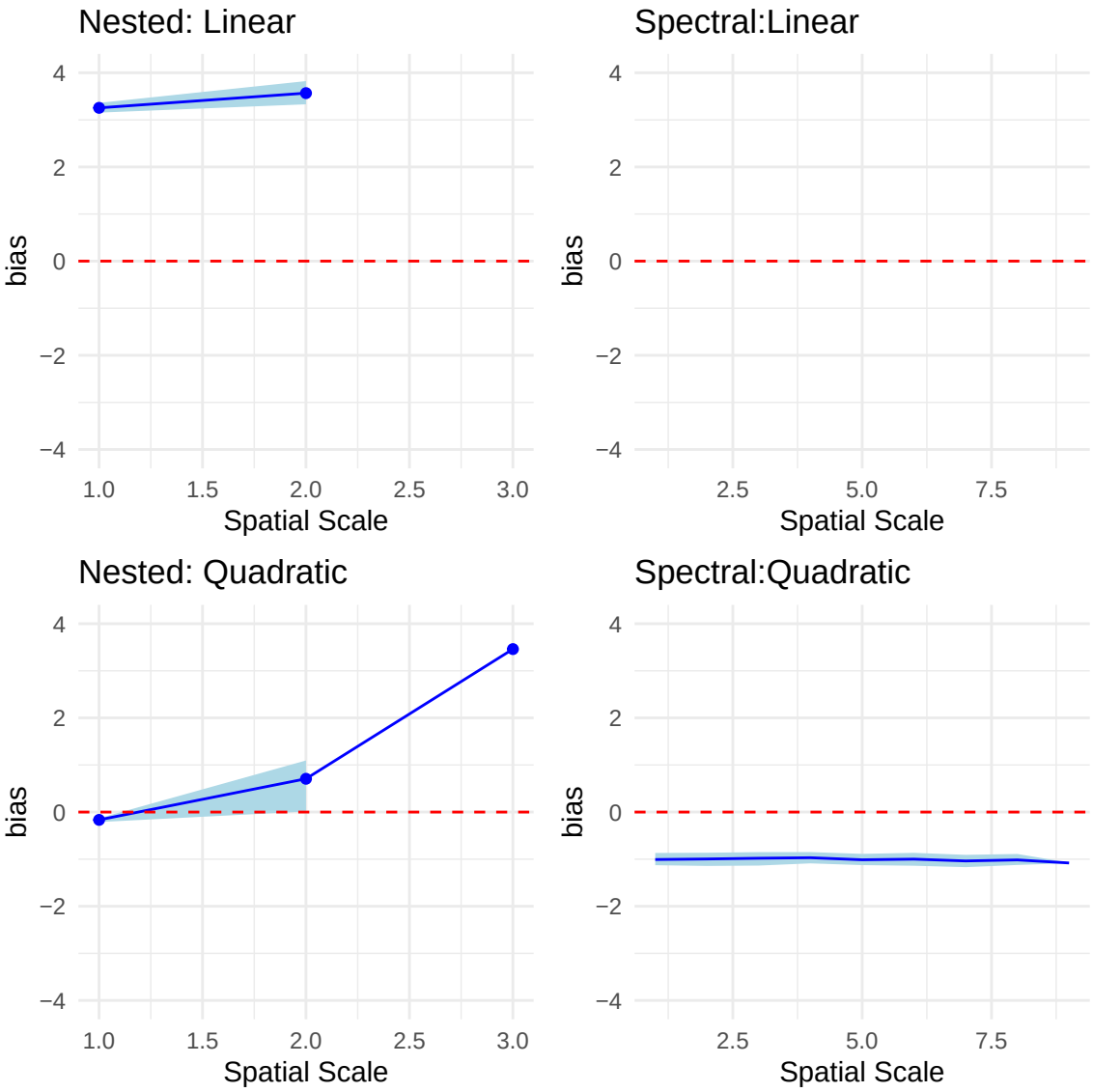


The bias of the estimates is the same as the results without an interaction... CHECK THIS

Simulation 6: Nonlinear Outcome Model

We repeat simulation 1 and 2 but the outcome model now includes a quadratic term of X. In particular, $Y = 2X + X^2 - Z + \epsilon$. The scale-specific analyses attempt to estimate both quadratic and linear coefficients of X. Neither method does a good job at recovering either of the coefficients.

Nested DGM



Spectral DGM