Algebraic Demonstration of How the Regression Estimator is Biased When Disturbances Covary

In order to see how the covariance between e_1 and e_3 leads to biased regression estimates, take a closer look at the regression formula for estimating b. The OLS estimator of b in equation (10.3) is:

$$\hat{b} = \frac{cov(Z,Y) - \frac{cov(Z,M)cov(Z,Y)}{var(M)}}{var(X) - \frac{cov(Z,M)cov(Z,M)}{var(M)}}. \tag{A.1}$$

Substituting for Y using equation (10.3) gives:

$$\hat{b} = \frac{Cov(Z,\alpha_3 + dZ + bM + e_3) - \frac{Cov(Z,M)Cov(M,\alpha_3 + dZ + bM + e_3)}{var(M)}}{Var(Z) - \frac{Cov(Z,M)Cov(Z,M)}{Var(M)}}$$

$$= \frac{dVar(Z) + bCov(Z,M) + Cov(Z,e_3) - \frac{Cov(Z,M)(dCov(Z,M) + bVar(M) + Cov(M,e_3))}{var(M)}}{var(Z) - \frac{cov(Z,M)cov(Z,M)}{var(M)}}. \tag{A.2}$$

Rearranging terms:

$$= b + \frac{Cov(Z,e_3) - \frac{Cov(M,e_3)}{Var(M)}}{var(Z) - \frac{Cov(Z,M)Cov(Z,M)}{Var(M)}}. \tag{A.3}$$

Substituting for M using equation (10.1) gives:

$$= b + \frac{Cov(Z,e_3) - \frac{Cov(\alpha_1 + aZ + e_1, e_3)}{Var(\alpha_1 + aZ + e_1)}}{Var(Z) - \frac{Cov(Z,\alpha_1 + aZ + e_1)Cov(Z,\alpha_1 + aZ + e_1)}{Var(\alpha_1 + aZ + e_1)}} = b + \frac{Cov(Z,e_3) - \frac{aCov(Z,e_3) + Cov(e_1,e_3)}{a^2Var(Z) + Var(e_1) + 2Cov(Z,e_1)}}{Var(Z) - \frac{(aVar(Z) + Cov(Z,e_1))^2}{a^2Var(Z) + Var(e_1) + 2Cov(Z,e_1)}}. \tag{A.4}$$

In order to see what this estimator would render if the sample were infinite, we evaluate the probability limit or "plim" of equation (10.8). Taking plims of both sides allows us to eliminate covariances between Z and the disturbances e_1 and e_3; in an infinite sample, these covariances become zero because Z is randomly assigned.

$$plim(\hat{b}) = b + \frac{\frac{Cov(e_1,e_3)}{a^2 Var(Z)+Var(e_1)}}{Var(Z) - \frac{(aVar(Z))^2}{a^2 Var(Z)+Var(e_1)}} = b + \frac{Cov(e_1,e_3)}{Var(e_1)}. \quad \text{(A.5)}$$

Equation (10.9) shows that this estimator is biased when $cov(e_1, e_3) \neq 0$.

The same approach can be used to show that the OLS estimator of d is biased as well:

$$plim(\hat{d}) = d - a\frac{Cov(e_1,e_3)}{Var(e_1)}. \quad \text{(A.6)}$$