Linear Model Simulations: Canonical IV DGP

Kellie Ottoboni

April 1, 2016

Suppose

$$
\begin{aligned}
Y_i &= \beta X_i + \epsilon_i \\
T_i &= \mathbb{I}(U_i + \delta_i > 0)
\end{aligned}
$$

where $(U_i, X_i, \epsilon_i, \delta_i)$ are IID multivariate normal with mean 0 across individuals $i = 1, \ldots, N$. For each i, we take the three objects X_i, U_i, and (ϵ_i, δ_i) to be mutually independent. Suppose further that

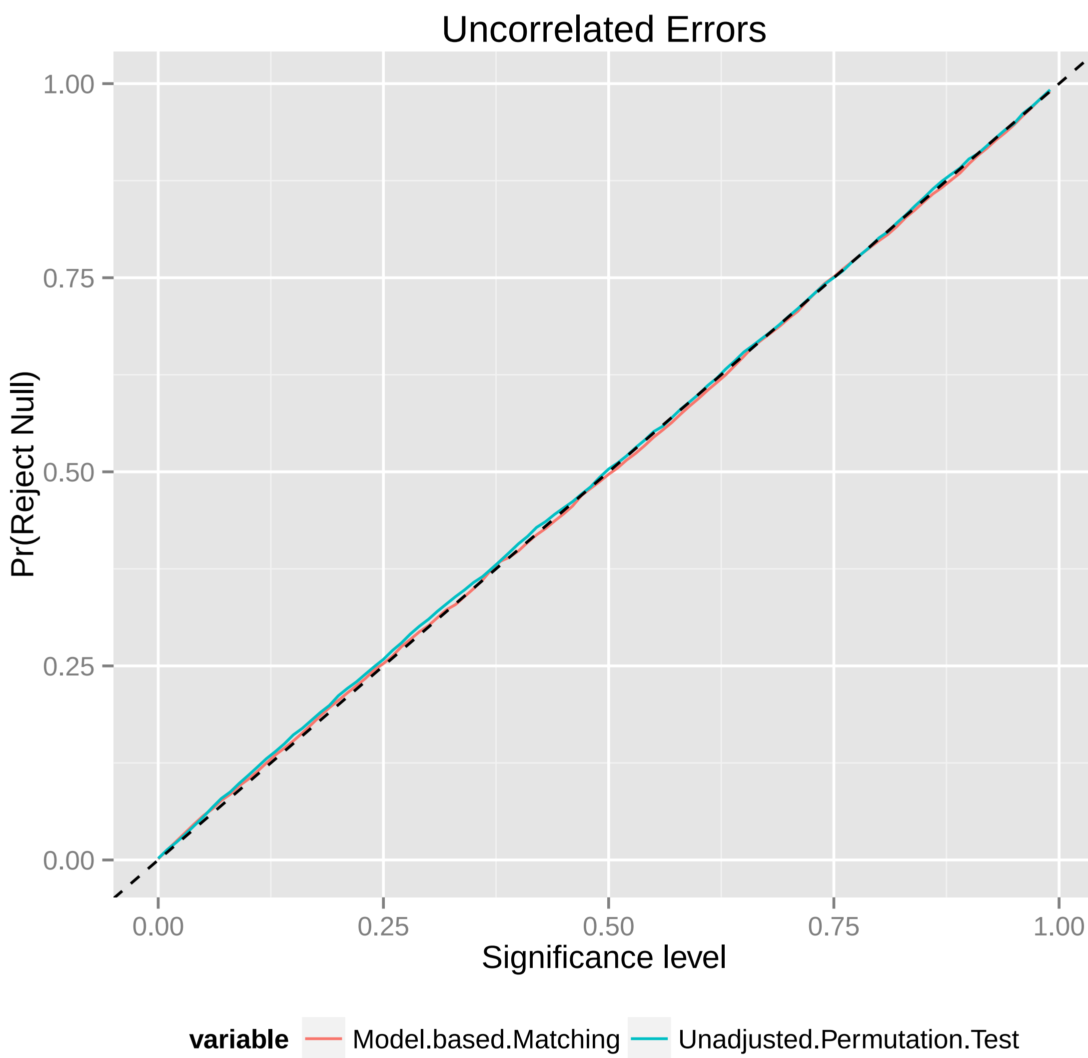
$$
\begin{aligned}
\mathrm{Var}(U_i) = \mathrm{Var}(X_i) &= 1 \\
\mathrm{Var}(\epsilon_i) = \mathrm{Var}(\delta_i) &= \sigma^2 \\
\mathrm{Cov}(\epsilon_i, \delta_i) &= \rho
\end{aligned}
$$

Estimate $\hat{Y}$ using OLS on X.

We have $N =$100 individuals. Suppose we use two strata, defined by whether $\hat{Y} \geq 0$ or $\hat{Y} < 0$.

We'll do the model-based matching permutation test using the difference in means as our statistic 5000 times to get a distribution of p-values under the null hypothesis of no treatment effect. The simulations are done **conditionally on X and ϵ**: we draw X and ϵ once to generate the potential outcomes Y, then draw random U and δ to generate treatment T for each simulation.

1 Uncorrelated errors

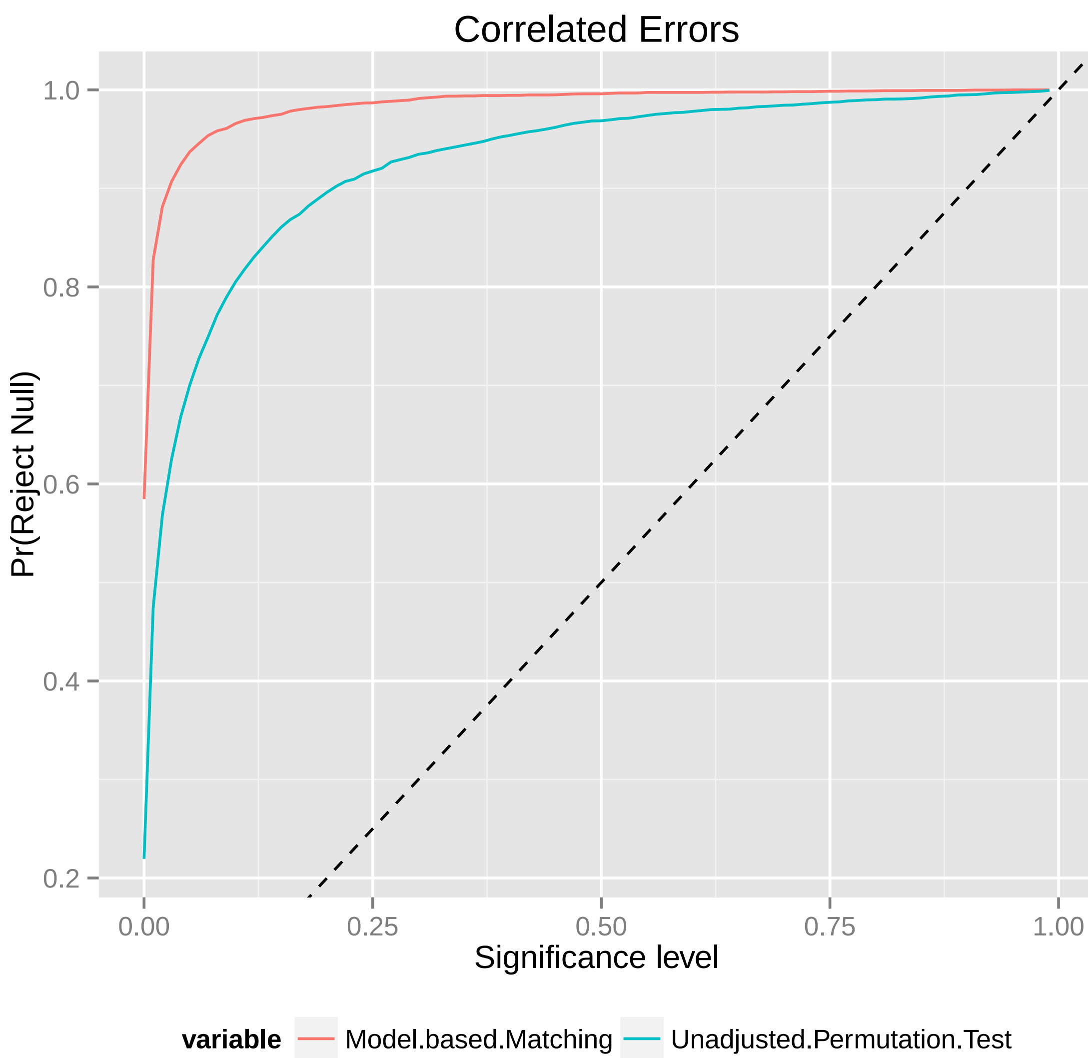


Unadjusted Permutation Test	0.055
Model-based Matching	0.057

Table 1: Proportion of tests rejected at level 0.05

2 Correlated errors

Here, we let $N = 100$, $\beta = 1$, $\sigma^2 = 1$, and $\rho = 3/4$.



Unadjusted Permutation Test	0.700
Model-based Matching	0.937

Table 2: Proportion of tests rejected at level 0.05