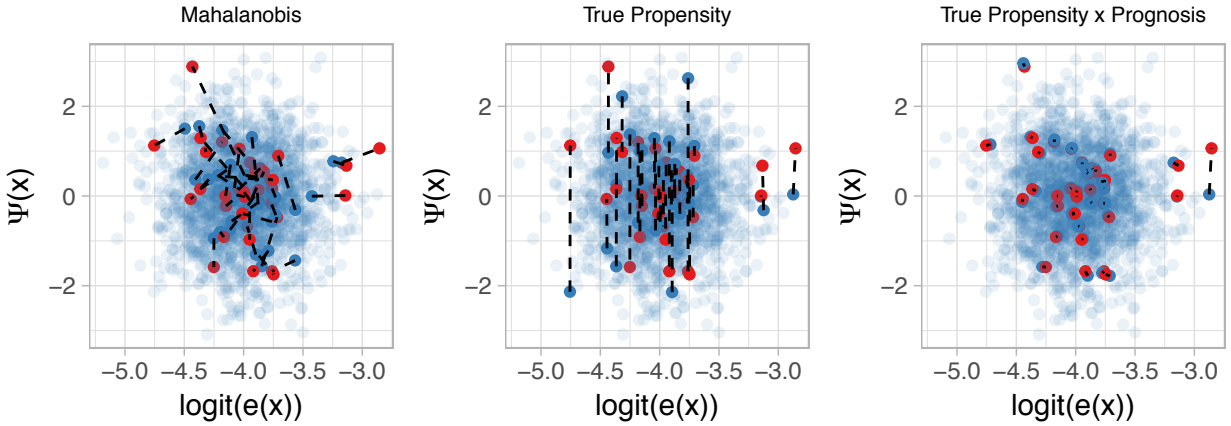
Replication of Dylan's Plots

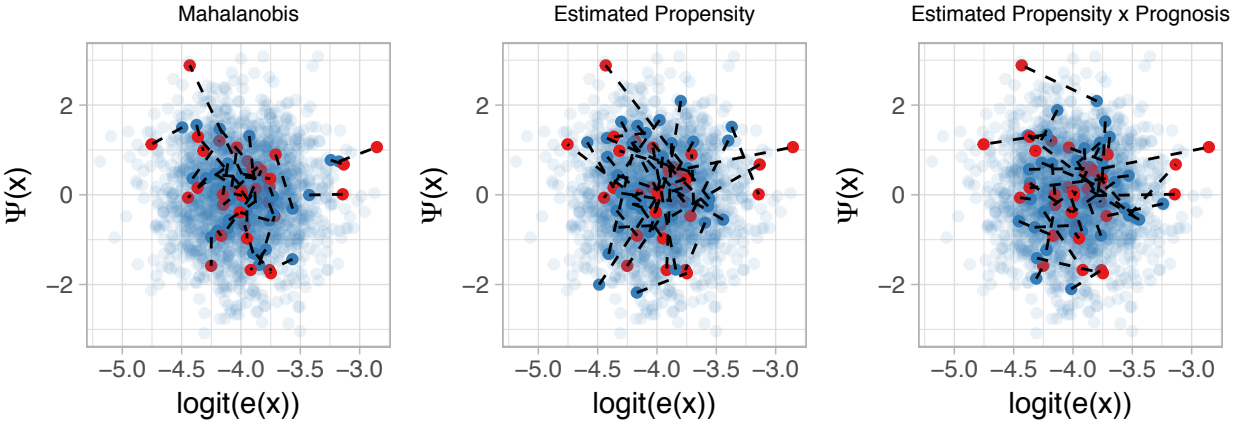
Basic Visualization

Here, I assume propensity and prognostic information are uncorrelated $\rho = 0$.

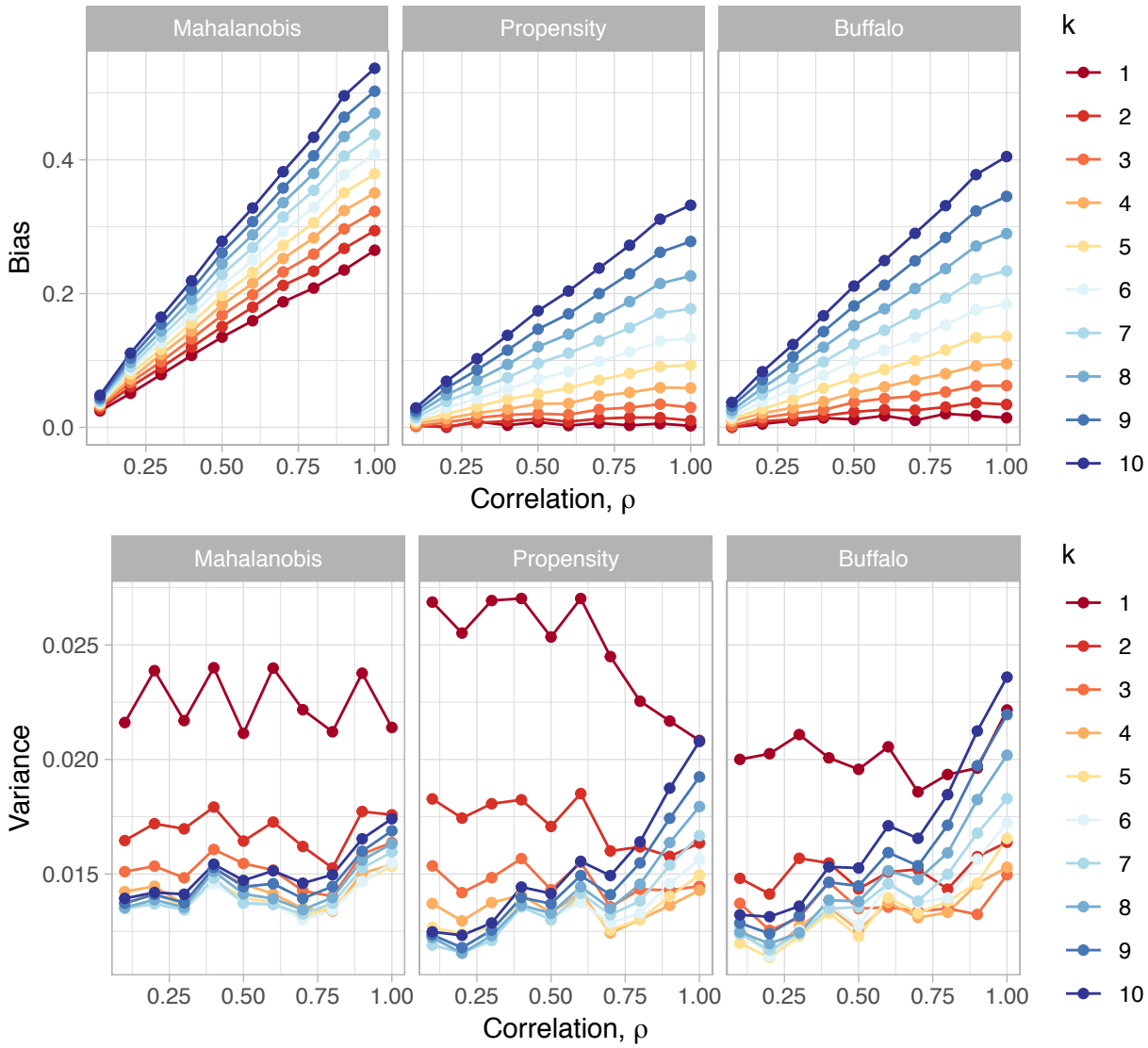
Below, we imagine that we know the true propensity and prognostic scores, and we match on those for buffalo and propensity score matching. Unsurprisingly, life is very good for buffalo matchers when the true scores are known.

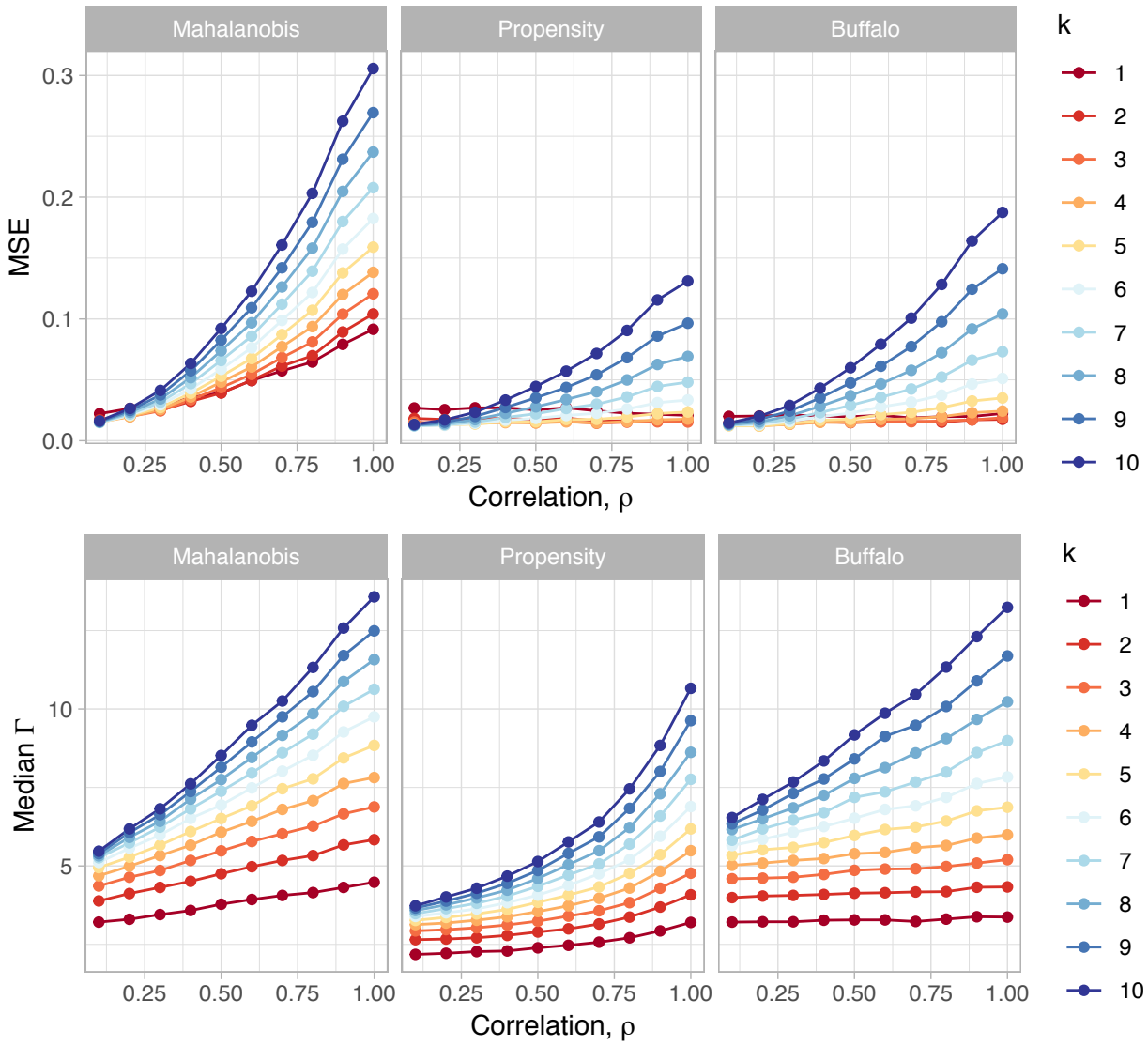


Now, we take the same data set and imagine that we don't know the true propensity and prognostic scores, and so must build them empirically. For propensity score matching, we build a logistic regression of treatment assigment on all of the variables, using the entire dataset. For buffalo, we use the propensity score logistic regression in addition to a prognostic score, which we fit on a subset of the controls that are chosen to be good matches to the treated individuals based on mahalanobis distance.



New color scheme




Mahalanobis
Propensity
Buffalo
MSE
0.0
0.1
0.2
0.3
0.25
0.50
0.75
1.00
Correlation, ρ
k
1
2
3
4
5
6
7
8
9
10
Median Γ
5
10

“‘