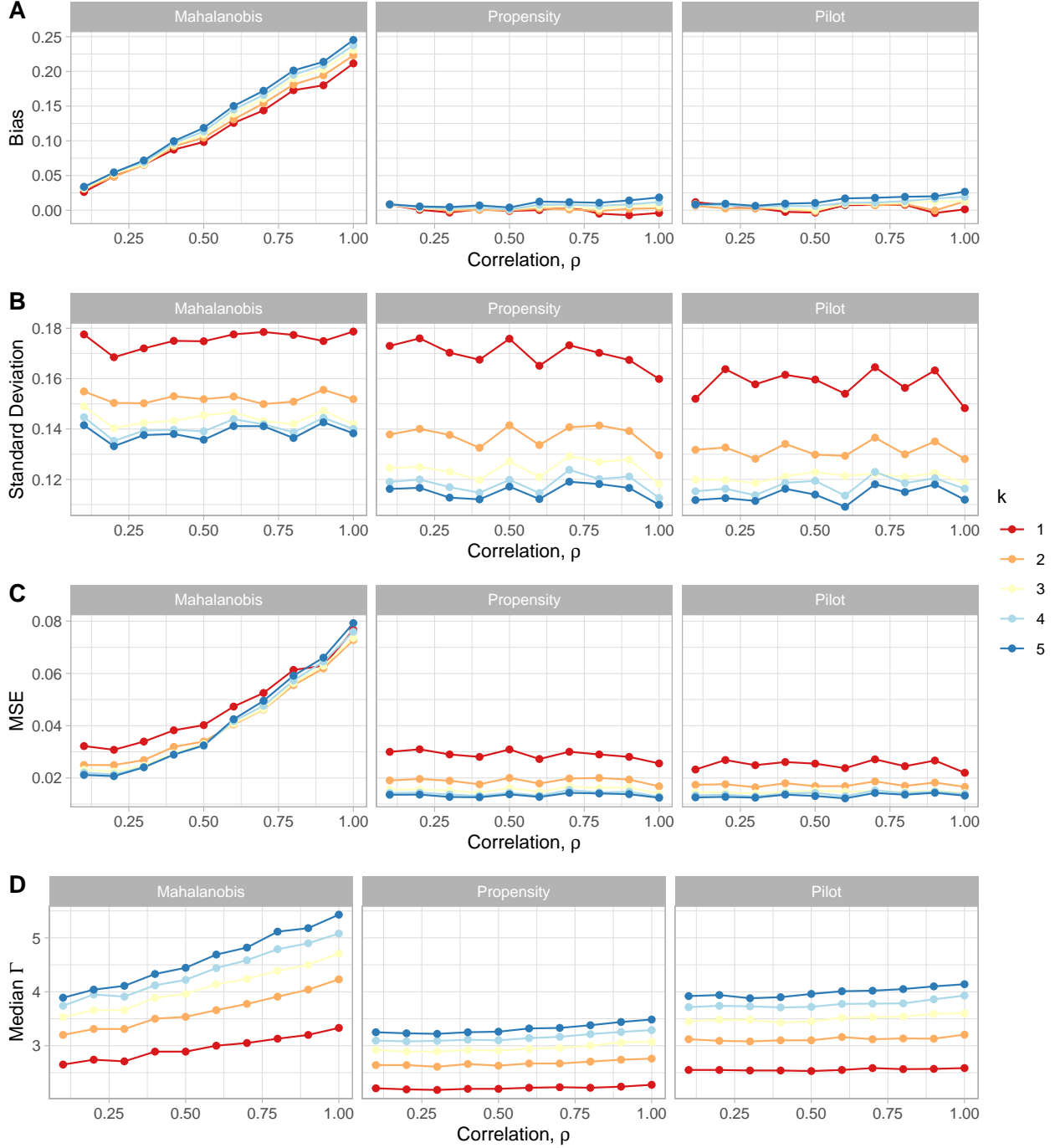


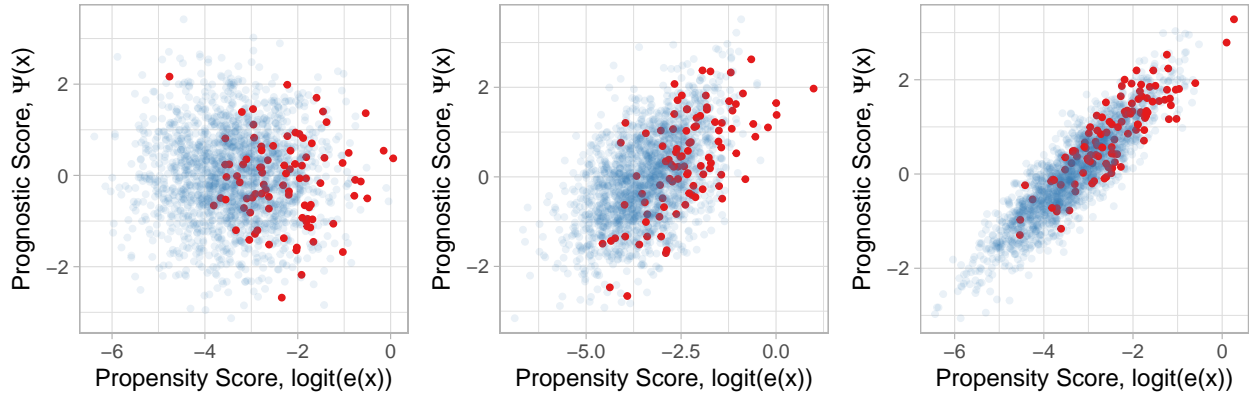
Supplementary Figures

Using the Prognostic Score to Reduce Heterogeneity in Observational Studies

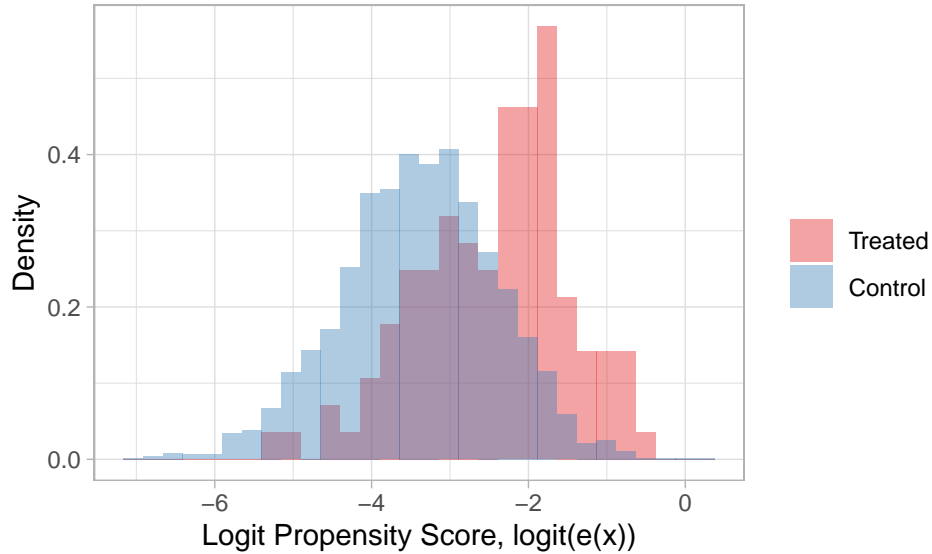
Rachael C. Aikens, Dylan Greaves, and Michael Baiocchi



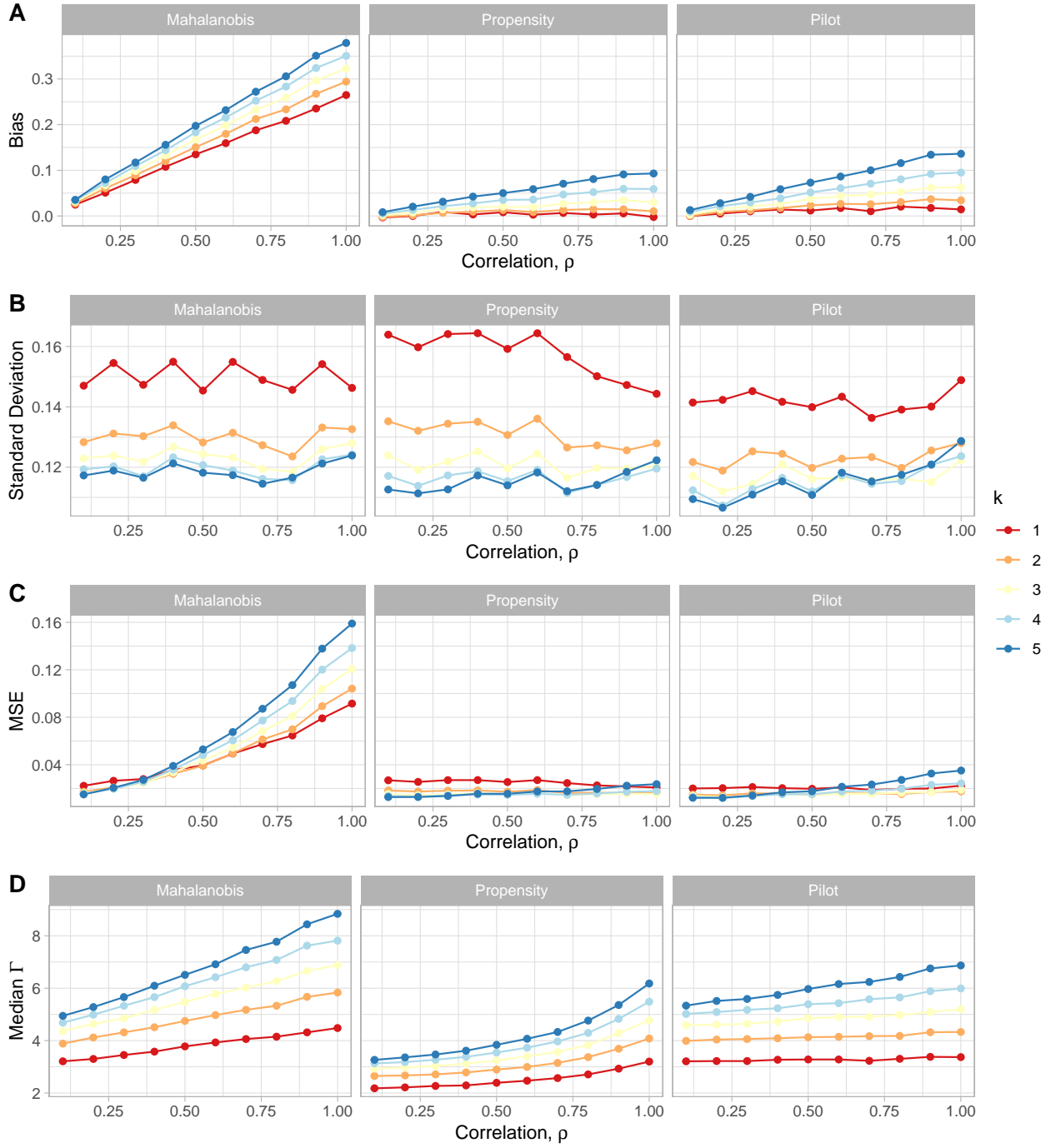
Supplementary Figure 1: Bias, standard deviation, MSE, and median Γ of matching estimators when the number of uninformative covariates is increased. All simulation parameters are the same as described in Section 4.2, except that the number of covariates, p , is increased to 50.



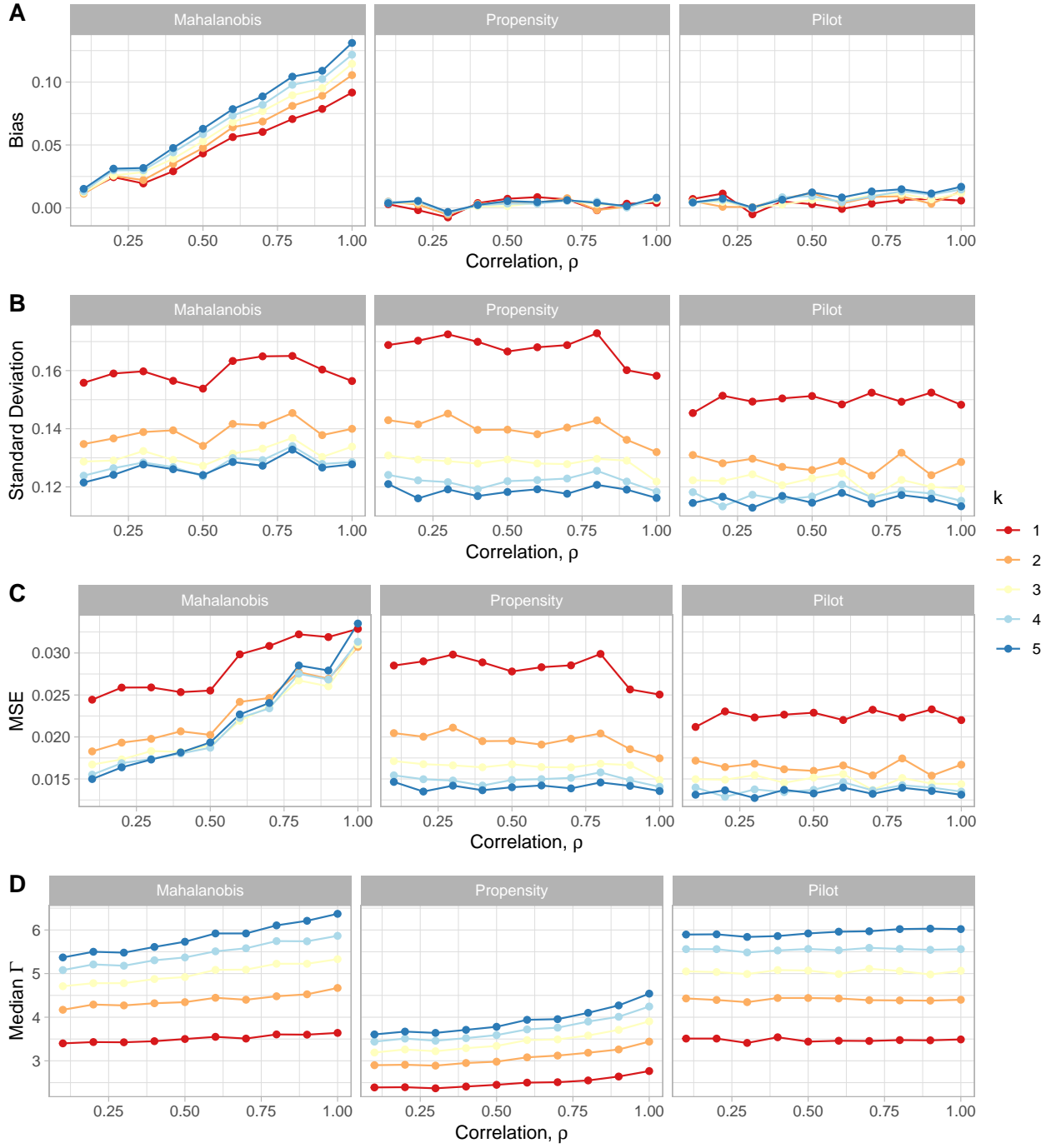
Supplementary Figure 2: Example Fisher-Mill plots for a scenario with diminished covariate overlap between treated and control individuals. Blue dots represent control individuals, red dots represent treated individuals, and dotted lines connect matched pairs. The correlation ρ between ϕ and Ψ is 0, 0.5, and 0.9, respectively in figures (A), (B) and (C). Simulations were carried out as described in Section 4.2, but with $\phi(X_i) = X_{i1} - 10/3$. This kept the sample size and number of treated individuals constant while increasing the separation of treated and control individuals in terms of the covariate X_1 .



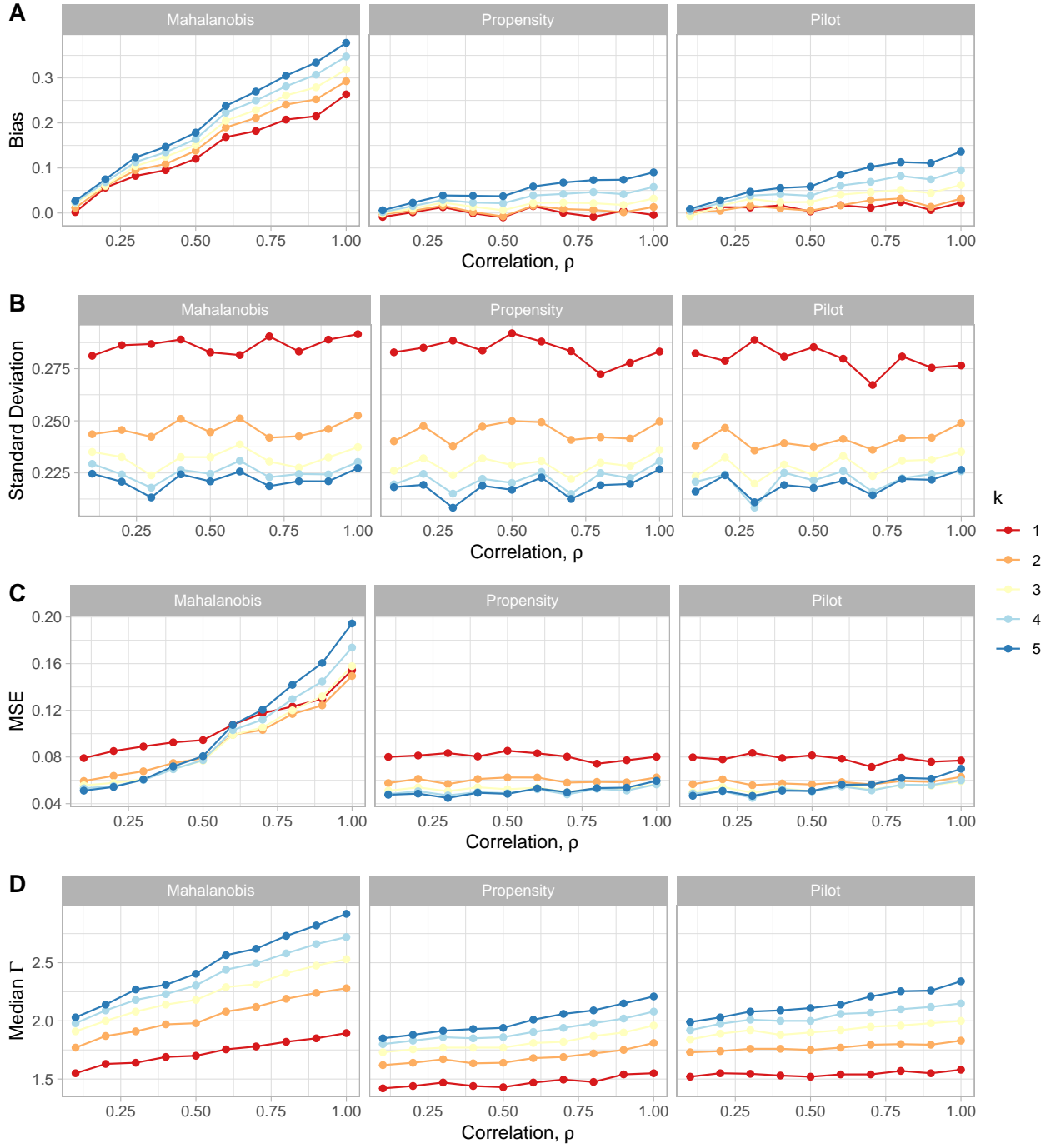
Supplementary Figure 3: Example overlaid density plot of logit propensity score in treated and control individuals for a simulation with diminished covariate overlap (in this simulation, $\rho = 0$). Note that there were approximately 19 control observations for every treated observation for each data set, so in practice the counts for the control densities are much larger.



Supplementary Figure 4: Bias, standard deviation, MSE, and median Γ from matching estimators for simulations with worse covariate overlap between treated and control observations.



Supplementary Figure 5: Bias, standard deviation, MSE, and median Γ from matching estimators for simulations with a heterogeneous treatment effect. Simulations were carried out as described in Section 4.2, but with the individual treatment effect determined by the covariates: $\tau(X_i) = 1 + X_i1/4$.



Supplementary Figure 6: Bias, standard deviation, MSE and median gamma design sensitivity when the random noise contributing to the outcome is increased. All simulation parameters are the same as described in Section 4.2, except that $\sigma = 2$. This increases the difficulty of fitting the prognostic score, diminishing the relative performance of pilot matching in terms of MSE and sensitivity.