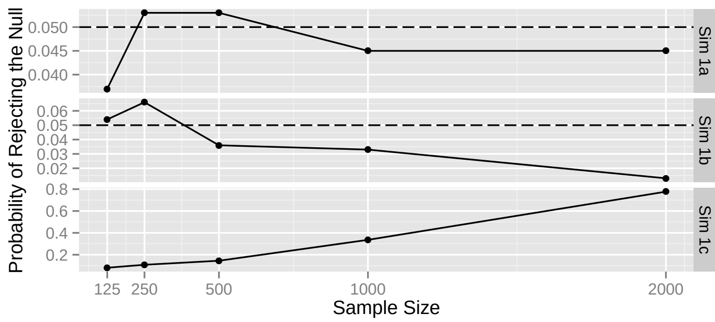

Figure 2: Empirical probability of rejecting the null when testing the null hypothesis that $\mu(1, W)$ is equal in distribution to $\mu(0, W)$ (Example 2) in Simulation 1. Table 1 indicates that the null is true in Simulations 1a and 1b, and the alternative is true in Simulation 1c.

expect. We are not aware of any other test devised for this hypothesis.

6.2 Simulation scenario 2: comparison with Racine et al. (2006)

We reproduced a simulation study from Section 4.1 of Racine et al. (2006) at sample size $n = 100$. In particular, we let $Y = 1 + \beta A(1 + W_2^2) + W_1 + W_2 + \epsilon$, where A, W_1, and W_2 are drawn independently from Bernoulli(0.5), Bernoulli(0.5), and $N(0, 1)$ distributions, respectively. The error term ϵ is unobserved and drawn from a $N(0, 1)$ distribution independently of all observed variables. The parameter β was varied over values $-0.5, -0.4, ..., 0.4, 0.5$ to achieve a range of distributions. The goal is to test whether $E_0(Y \mid A, W) = E_0(Y \mid W)$ almost surely, or equivalently, that $\mu(1, W) - \mu(0, W) = 0$ almost surely.

Due to computational constraints, we only ran the 'Bootstrap I test' to evaluate significance of the method of Racine et al. (2006). As the authors report, this method is anticonservative relative to their 'Bootstrap II test' and indeed achieves lower power (but proper type I error control) in their simulations.

Except for two minor modifications, our implementation of the method in Example 1 is similar to that as for Simulation 1. For a fair comparison with Racine et al. (2006), in this simulation study, we estimated $P_0(A = 1 \mid W)$ rather than treating it as known. We did this using the same Super Learner library and the '`family=binomial`' setting to account for