Supplement Material for "Consistent Tests for Conditional Treatment Effects"

Yu-Chin Hsu

Institute of Economics

Academia Sinica

A Monte-Carlo Simulations

In this section, we conduct small-scale Monte-Carlo simulations to illustrate the finite sample performance of our test.

A.1 Simulation Designs

Recall that $\Lambda(a) = \exp(a)/(1 + \exp(a))$ is the logistic CDF. For the finite-sample size properties, we consider the following data generating processes (DGP):

- **DGP1:** Let the DGP1 be:

$$
\begin{aligned}
&X_1 = U_1, \quad X_2 = U_2, \quad D = 1(U_d < \Lambda(X_1 + X_2 - 1)), \\
&Y(1) = 0.2(X_1 - 0.5)(X_2 - 0.5) + U_y, \\
&Y(0) = 0, \\
&Y = DY(1) + (1 - D)Y(0),
\end{aligned}
$$

  where U_1, U_2 and U_d are uniform distributions over $[0, 1]$ and U_y is normal distribution with mean zero and variance 0.25^2. U_1, U_2, U_d and U_y are mutually independent.

- **DGP2:** Let DGP2 be the same as DGP1 except that $Y(1) = 0.2(X_1 - 0.5)X_2 \cdot 1(X_1 > 0.5) + U_y$.

- **DGP3:** Let the DGP3 be the same as DGP1 except that $D = 1(U_d < 0.3 + 0.2(X_1 + X_2))$.

- **DGP4:** Let the DGP4 be the same as DGP2 except that $D = 1(U_d < 0.3 + 0.2(X_1 + X_2))$.

- **DGP5:** Let the DGP5 be:

$$
\begin{aligned}
&X_1 = U_1, \quad X_2 = U_2, \quad X_3 = U_1, \quad X_4 = U_2, \\
&D = 1(U_d < \Lambda(0.5(X_1 + X_2 + X_3 + X_4 - 2))), \\
&Y(1) = 0.8(X_1 - 0.5)(X_2 - 0.5)(X_3 - 0.5)(X_4 - 0.5) + U_y, \\
&Y(0) = 0, \\
&Y = DY(1) + (1 - D)Y(0),
\end{aligned}
$$

  where U_1, U_2, U_3, U_4 and U_d are uniform distributions over $[0, 1]$ and U_y is normal distribution with mean zero and variance 0.25^2. U_1, U_2, U_3, U_4, U_d and U_y are mutually independent.

- **DGP6:** Let DGP6 be the same as DGP5 except that $Y(1) = 0.8(X_1 - 0.5)X_2X_3X_4 \cdot 1(X_1 > 0.5) + U_y$.
- **DGP7:** Let DGP7 be the same as DGP5 except that $D = 1(U_d < 0.3 + 0.1(X_1 + X_2 + X_3 + X_4))$.
- **DGP8:** Let DGP8 be the same as DGP6 except that $D = 1(U_d < 0.3 + 0.1(X_1 + X_2 + X_3 + X_4))$.

For DGPs 1, 3, 5 and 7, $CATE(x_1) = 0$ for all $x_1 \in [0, 1]$ and the null hypothesis holds. We use these examples to illustrate the size properties of our test when the null hypothesis is the least favorable case. For DGPs 2, 4, 6, 8, $CATE(X_1) = 0.1X_1 - 0.05 > 0$ when $X_1 > 0.5$ and $CATE(X_1) = 0$ when $X_1 \leq 0.5$. We use these examples to illustrate the size properties of our test when the null hypothesis is not the least favorable case.

For the finite-sample power properties, we consider the following DGPs:

- **DGP9:** Let DGP9 be the same as DGP1 except that $Y(1) = 0.2(X_1 - 0.5)X_2 + U_y$.
- **DGP10:** Let DGP10 be the same as DGP1 except that $Y(1) = 0.2(X_1 - 0.5)X_2 \cdot 1(X_1 \leq 0.5) + U_y$.
- **DGP11:** Let DGP2 be the same as DGP9 except that $D = 1(U_d < 0.3 + 0.2(X_1 + X_2))$.
- **DGP12:** Let DGP2 be the same as DGP10 except that $D = 1(U_d < 0.3 + 0.2(X_1 + X_2))$.
- **DGP13:** Let DGP13 be the same as DGP5 except that $Y(1) = 0.8(X_1 - 0.5)X_2X_3X_4 + U_y$.
- **DGP14:** Let the DGP14 be the same as DGP5 except that $Y(1) = 0.8(X_1 - 0.5)X_2X_3X_4 \cdot 1(X_1 \leq 0.5) + U_y$.
- **DGP15:** Let DGP15 be the same as DGP13 except that $D = 1(U_d < 0.3 + 0.1(X_1 + X_2 + X_3 + X_4))$.
- **DGP16:** Let DGP16 be the same as DGP14 except that $D = 1(U_d < 0.3 + 0.1(X_1 + X_2 + X_3 + X_4))$.

For DGPs 9-16, we have $CATE(x_1) = 0.5x_1 < 0$ when $x_1 < 0.5$. We use these example to illustrate the power properties of our test.

A.2 Implementation Details

For each case, we consider sample sizes $N = 500$, 1,000 and 2,000. For IV functions, we consider

$$\begin{aligned}
\mathcal{G} &= \{g_\ell(\cdot) = 1(\cdot \in C_\ell) : \ell \equiv (x_s, r) \in \mathcal{L}\}, \text{ where} \\
C_{\ell_s} &= \left(\times_{j=1}^{d_s}[x_j, x_j + r]\right) \cap \mathcal{X}_s \text{ and} \\
\mathcal{L} &= \left\{(x_s, q^{-1}) : \ q \cdot x_s \in \{0, 1, 2, \cdots, q-1\}^{d_s}, \text{ and } q = 1, 2, \cdots, q_1\right\},
\end{aligned} \tag{A.1}$$

where q_1 is a natural number. For $N = 500$ and 1000, we consider $q_1 = 5$ and 10, and for $N = 2000$, we consider $q_1 = 10$ and 15. Regarding the power series in the SLE, for cases with 2 covriates, we consider all the power series up to order 2 and 3 when $n = 500$ and 1,000 and we consider all the power series up to order 3 and 4 when $N = 2,000$. for cases with 4 covariates, we consider all the power series up to order 1 and 2 when $N = 500$ and 1,000 and we consider all the power series up to order 2 and 3 when $N = 2,000$. We set $a_N = -(0.3\ln(N))^{1/2}$, $B_N = (0.4\ln(N)/\ln\ln(N))^{1/2}$ and $\eta = 10^{-6}$ following AS's suggestions. The rejection rates are approximated by 1,000 repetitions and the simulated critical values are approximated by 1,000 simulations. The significance level is set at 5%.

A.3 Simulation Results

The finite-sample size and power results are summarized in Tables 1 and 2, respectively. For the size cases, we can see that our test can control the size well over all cases and the size is not sensitive to the choices of power series for the SLE and the choices of the IV functions. For DGPs 1, 3, 5 and 7, the null hypotheses are at the LFC cases and the sizes are close to 5%, the nominal level. For DGPs 2, 4, 6 and 8, the null is not at the LFC cases, we see the sizes are under the nominal level. In general, we will expect the finite sample size will increase to 5% when sample size increases due to the GMS method we use in the simulations. We do not see this in these cases because under the set up, we have $CATE(X_1) = 0.1X_1 - 0.05 > 0$ when $X_1 > 0.5$ which is very close to zero. Therefore, the GMS is not very effective even if sample size is 2,000 in our examples. For the power properties, we can see that the rejection rates are all above the 5% nominal level and the rejection rate increases when the sample size increases. This confirms that our test is consistent.

Table 1: Size Properties

N	order	IV (q_1)	DGP1	DGP2	DGP3	DGP4
500	2	5	0.046	0.018	0.053	0.020
500	2	10	0.044	0.014	0.064	0.023
500	3	5	0.054	0.018	0.042	0.026
500	3	10	0.049	0.020	0.053	0.019
1,000	2	5	0.047	0.017	0.054	0.012
1,000	2	10	0.055	0.028	0.047	0.013
1,000	3	5	0.055	0.016	0.048	0.011
1,000	3	10	0.050	0.014	0.057	0.014
2,000	3	10	0.054	0.006	0.049	0.004
2,000	3	15	0.045	0.017	0.043	0.012
2,000	4	10	0.054	0.013	0.049	0.010
2,000	4	15	0.042	0.017	0.048	0.011
N	order	IV (q_1)	DGP5	DGP6	DGP7	DGP8
500	1	5	0.039	0.015	0.050	0.015
500	1	10	0.046	0.009	0.043	0.014
500	2	5	0.045	0.015	0.061	0.012
500	2	10	0.056	0.019	0.043	0.015
1,000	1	5	0.042	0.007	0.052	0.013
1,000	1	10	0.038	0.009	0.046	0.009
1,000	2	5	0.037	0.016	0.052	0.013
1,000	2	10	0.063	0.013	0.050	0.009
2,000	2	10	0.052	0.011	0.053	0.012
2,000	2	15	0.054	0.017	0.054	0.007
2,000	3	10	0.059	0.012	0.055	0.014
2,000	3	15	0.053	0.016	0.053	0.011

Significance level is set at 5%.

Table 2: Power Properties

N	order	IV (q_1)	DGP9	DGP10	DGP11	DGP12
500	2	5	0.110	0.207	0.103	0.212
500	2	10	0.112	0.198	0.121	0.206
500	3	5	0.153	0.235	0.154	0.239
500	3	10	0.133	0.231	0.133	0.230
1,000	2	5	0.247	0.348	0.241	0.343
1,000	2	10	0.224	0.319	0.231	0.322
1,000	3	5	0.214	0.367	0.204	0.349
1,000	3	10	0.239	0.324	0.230	0.320
2,000	3	10	0.447	0.575	0.459	0.579
2,000	3	15	0.439	0.565	0.434	0.571
2,000	4	10	0.444	0.568	0.460	0.572
2,000	4	15	0.437	0.585	0.454	0.587
N	order	IV (q_1)	DGP13	DGP14	DGP15	DGP16
500	1	5	0.106	0.205	0.143	0.215
500	1	10	0.106	0.204	0.142	0.214
500	2	5	0.136	0.218	0.131	0.209
500	2	10	0.149	0.239	0.134	0.230
1,000	1	5	0.243	0.354	0.226	0.364
1,000	1	10	0.217	0.330	0.227	0.355
1,000	2	5	0.223	0.344	0.234	0.344
1,000	2	10	0.242	0.326	0.237	0.364
2,000	2	10	0.467	0.579	0.466	0.562
2,000	2	15	0.451	0.580	0.461	0.592
2,000	3	10	0.463	0.566	0.448	0.590
2,000	3	15	0.441	0.532	0.444	0.598

Significance level is set at 5%.