

Table 1: Experimental results on BC(homo) Dataset. The top result is highlighted in bold, and the runner-up is underlined. / means not applicable.

| Methods | $\epsilon_{average}$ | | | | | | $\epsilon_{individual}$ | | | | | |
|----------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| | Within Sample | | | Out-of Sample | | | Within Sample | | | Out-of Sample | | |
| | AME | ASE | ATE | AME | ASE | ATE | IME | ISE | ITE | IME | ISE | ITE |
| TARNET+z | 0.1573 ± 0.0045 | 0.0824 ± 0.0149 | 0.2046 ± 0.0193 | 0.1492 ± 0.0379 | 0.0855 ± 0.0147 | 0.1982 ± 0.0380 | 0.2006 ± 0.0250 | 0.1161 ± 0.0159 | 0.2444 ± 0.0239 | 0.2809 ± 0.0507 | 0.1209 ± 0.0152 | 0.3062 ± 0.0627 |
| CFR+z | <u>0.0788</u> ± 0.0096 | 0.1157 ± 0.0076 | 0.2323 ± 0.0196 | 0.0770 ± 0.0099 | 0.1157 ± 0.0075 | 0.2306 ± 0.0196 | <u>0.0796</u> ± 0.0091 | 0.1158 ± 0.0076 | 0.2325 ± 0.0196 | <u>0.1000</u> ± 0.0388 | 0.1157 ± 0.0075 | 0.2405 ± 0.0166 |
| GEst | 0.1872 ± 0.0672 | 0.2369 ± 0.0607 | 0.1422 ± 0.0562 | 0.1953 ± 0.0611 | 0.2391 ± 0.0617 | 0.1302 ± 0.0524 | 0.2307 ± 0.0493 | 0.2603 ± 0.0586 | 0.1877 ± 0.0495 | 0.2388 ± 0.0431 | 0.2623 ± 0.0592 | 0.1790 ± 0.0440 |
| ND+z | 0.2375 ± 0.0450 | 0.0316 ± 0.0104 | 0.0790 ± 0.0236 | 0.2380 ± 0.0458 | 0.0323 ± 0.0122 | 0.0768 ± 0.0254 | 0.2377 ± 0.0448 | <u>0.0321</u> ± 0.0101 | 0.0792 ± 0.0236 | 0.2477 ± 0.0400 | <u>0.0379</u> ± 0.0099 | 0.1068 ± 0.0172 |
| NetEst | 0.1059 ± 0.0609 | <u>0.0281</u> ± 0.0297 | <u>0.0287</u> ± 0.0288 | 0.0987 ± 0.0603 | <u>0.0257</u> ± 0.0276 | 0.0356 ± 0.0288 | 0.1366 ± 0.0481 | 0.0631 ± 0.0205 | 0.0994 ± 0.0214 | 0.1680 ± 0.0620 | 0.0920 ± 0.0316 | 0.1507 ± 0.0647 |
| TNet | 0.1045 ± 0.0610 | 0.0502 ± 0.0559 | 0.0473 ± 0.0220 | 0.1045 ± 0.0610 | 0.0502 ± 0.0559 | 0.0473 ± 0.0220 | 0.1045 ± 0.0610 | 0.0502 ± 0.0559 | <u>0.0473</u> ± 0.0220 | 0.1045 ± 0.0610 | 0.0502 ± 0.0559 | 0.0473 ± 0.0220 |
| RRNet | 0.0884 ± 0.0495 | 0.0445 ± 0.0158 | 0.0768 ± 0.0219 | 0.0892 ± 0.0505 | 0.0447 ± 0.0160 | 0.0782 ± 0.0234 | 0.0915 ± 0.0462 | 0.0452 ± 0.0154 | 0.0865 ± 0.0260 | 0.0917 ± 0.0491 | 0.0453 ± 0.0159 | 0.0887 ± 0.0282 |
| SPNet+z | 0.0802 ± 0.0632 | 0.0446 ± 0.0155 | 0.0570 ± 0.0188 | 0.0931 ± 0.0607 | 0.0394 ± 0.0215 | 0.0385 ± 0.0183 | 0.1153 ± 0.0482 | 0.0741 ± 0.0166 | 0.1094 ± 0.0274 | 0.1358 ± 0.0566 | 0.0604 ± 0.0145 | 0.0953 ± 0.0342 |
| Ours | 0.0532 ± 0.0500 | 0.0253 ± 0.0206 | 0.0377 ± 0.0245 | 0.0557 ± 0.0535 | 0.0259 ± 0.0273 | <u>0.0281</u> ± 0.0218 | 0.0542 ± 0.0502 | 0.0257 ± 0.0206 | 0.0431 ± 0.0196 | 0.0561 ± 0.0535 | 0.0260 ± 0.0273 | 0.0400 ± 0.0208 |

Table 2: Experimental results on BC(hete) Dataset. The top result is highlighted in bold, and the runner-up is underlined. / means not applicable.

| Methods | $\epsilon_{average}$ | | | | | | $\epsilon_{individual}$ | | | | | |
|----------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| | Within Sample | | | Out-of Sample | | | Within Sample | | | Out-of Sample | | |
| | AME | ASE | ATE | AME | ASE | ATE | IME | ISE | ITE | IME | ISE | ITE |
| TARNET+z | 0.2538 ± 0.1127 | 0.1657 ± 0.0563 | 0.3866 ± 0.0711 | 0.2619 ± 0.1054 | 0.1701 ± 0.0594 | 0.4044 ± 0.1334 | 0.3455 ± 0.0654 | 0.2122 ± 0.0528 | 0.4695 ± 0.0720 | 0.5500 ± 0.2643 | 0.2207 ± 0.0510 | 0.6349 ± 0.3280 |
| CFR+z | 0.1580 ± 0.0189 | 0.2071 ± 0.0237 | 0.4067 ± 0.0407 | 0.1559 ± 0.0203 | 0.2076 ± 0.0245 | 0.4061 ± 0.0449 | 0.1825 ± 0.0129 | 0.2092 ± 0.0233 | 0.4316 ± 0.0351 | 0.2058 ± 0.0432 | 0.2098 ± 0.0242 | 0.4422 ± 0.0456 |
| GEst | 0.2734 ± 0.1240 | 0.4257 ± 0.0973 | 0.2916 ± 0.1119 | 0.2722 ± 0.1308 | 0.4277 ± 0.1012 | 0.2873 ± 0.1230 | 0.3334 ± 0.1082 | 0.4592 ± 0.0934 | 0.3546 ± 0.0947 | 0.3832 ± 0.1474 | 0.4612 ± 0.0972 | 0.3958 ± 0.1486 |
| ND+z | 0.4124 ± 0.0702 | 0.0451 ± 0.0201 | 0.1330 ± 0.0205 | 0.4111 ± 0.0737 | 0.0486 ± 0.0206 | 0.1326 ± 0.0261 | 0.4226 ± 0.0673 | <u>0.0562</u> ± 0.0146 | 0.1941 ± 0.0246 | 0.4330 ± 0.0602 | 0.0666 ± 0.0137 | 0.2211 ± 0.0321 |
| NetEst | 0.1643 ± 0.1337 | 0.0450 ± 0.0180 | 0.0594 ± 0.0262 | 0.1857 ± 0.1168 | <u>0.0405</u> ± 0.0252 | 0.0343 ± 0.0179 | 0.2199 ± 0.1039 | 0.0667 ± 0.0171 | 0.1731 ± 0.0368 | 1.5505 ± 2.5329 | 1.1347 ± 1.9028 | 1.0924 ± 1.7278 |
| TNet | 0.1216 ± 0.0864 | 0.0537 ± 0.0524 | <u>0.0429</u> ± 0.0301 | 0.1257 ± 0.0737 | 0.0537 ± 0.0511 | 0.0481 ± 0.0289 | <u>0.1731</u> ± 0.0459 | 0.0655 ± 0.0465 | 0.1458 ± 0.0175 | 0.1915 ± 0.0542 | 0.0650 ± 0.0458 | <u>0.1740</u> ± 0.0621 |
| RRNet | 0.1628 ± 0.1183 | <u>0.0476</u> ± 0.0416 | 0.0919 ± 0.0581 | 0.1636 ± 0.1162 | 0.0474 ± 0.0424 | 0.0961 ± 0.0640 | 0.1967 ± 0.0979 | 0.0591 ± 0.0364 | 0.1719 ± 0.0485 | 0.1974 ± 0.0969 | <u>0.0595</u> ± 0.0365 | 0.1764 ± 0.0595 |
| SPNet+z | <u>0.0706</u> ± 0.0738 | 0.0714 ± 0.0256 | 0.1169 ± 0.0317 | <u>0.0805</u> ± 0.0716 | 0.0671 ± 0.0230 | 0.0841 ± 0.0419 | 0.1743 ± 0.0338 | 0.1054 ± 0.0138 | 0.2299 ± 0.0288 | <u>0.1901</u> ± 0.0480 | 0.0906 ± 0.0190 | 0.2012 ± 0.0480 |
| Ours | 0.0434 ± 0.0213 | 0.0337 ± 0.0160 | 0.0370 ± 0.0292 | 0.0380 ± 0.0206 | 0.0320 ± 0.0172 | <u>0.0401</u> ± 0.0212 | 0.1033 ± 0.0136 | 0.0465 ± 0.0095 | <u>0.1470</u> ± 0.0272 | 0.0976 ± 0.0181 | 0.0455 ± 0.0096 | 0.1475 ± 0.0259 |

Table 3: Experimental results on Flickr(homo) Dataset. The top result is highlighted in bold, and the runner-up is underlined. / means not applicable.

| Methods | $\epsilon_{average}$ | | | | | | $\epsilon_{individual}$ | | | | | |
|----------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| | Within Sample | | | Out-of Sample | | | Within Sample | | | Out-of Sample | | |
| | AME | ASE | ATE | AME | ASE | ATE | IME | ISE | ITE | IME | ISE | ITE |
| TARNET+z | 0.0783 ± 0.0418 | 0.0874 ± 0.0213 | 0.2025 ± 0.0396 | 0.0976 ± 0.0506 | 0.0724 ± 0.0184 | 0.1356 ± 0.0567 | 0.1362 ± 0.0254 | 0.1103 ± 0.0194 | 0.2358 ± 0.0383 | 1.0869 ± 1.2258 | 0.1011 ± 0.0185 | 1.0889 ± 1.2270 |
| CFR+z | 0.0579 ± 0.0247 | 0.0785 ± 0.0070 | 0.1651 ± 0.0121 | 0.0507 ± 0.0192 | 0.0783 ± 0.0070 | 0.1581 ± 0.0097 | 0.0509 ± 0.0240 | 0.0786 ± 0.0070 | 0.1654 ± 0.0120 | 0.3465 ± 0.4635 | 0.0786 ± 0.0069 | 0.4102 ± 0.4278 |
| GEat | 0.1551 ± 0.0130 | 0.2475 ± 0.0676 | 0.0805 ± 0.0325 | 0.1511 ± 0.0137 | 0.2494 ± 0.0470 | 0.0805 ± 0.0278 | 0.1779 ± 0.0122 | 0.2656 ± 0.0378 | 0.1268 ± 0.0160 | 0.2867 ± 0.2172 | 0.2677 ± 0.0372 | 0.2471 ± 0.2352 |
| ND+z | 0.1416 ± 0.0240 | 0.0204 ± 0.0093 | 0.0478 ± 0.0216 | 0.1435 ± 0.0364 | 0.0226 ± 0.0101 | 0.0485 ± 0.0236 | 0.1427 ± 0.0246 | 0.0221 ± 0.0090 | 0.0501 ± 0.0178 | 0.3849 ± 0.2395 | 0.0348 ± 0.0078 | 0.3453 ± 0.2772 |
| NetEst | 0.0515 ± 0.0520 | 0.0355 ± 0.0317 | 0.0715 ± 0.0381 | 0.0470 ± 0.0590 | 0.0338 ± 0.0330 | 0.0529 ± 0.0395 | 0.0844 ± 0.0406 | 0.0566 ± 0.0253 | 0.1043 ± 0.0312 | 0.2934 ± 0.3001 | 0.2809 ± 0.3307 | 0.3068 ± 0.1960 |
| TNet | 0.0319 ± 0.0248 | 0.0274 ± 0.0089 | 0.0735 ± 0.0240 | 0.0299 ± 0.0231 | 0.0277 ± 0.0113 | 0.0715 ± 0.0214 | 0.0347 ± 0.0282 | 0.0276 ± 0.0313 | 0.0752 ± 0.0263 | 0.0561 ± 0.0648 | 0.0286 ± 0.0331 | 0.0918 ± 0.0555 |
| RRNet | <u>0.0296</u> ± 0.0123 | <u>0.0251</u> ± 0.0172 | 0.0199 ± 0.0179 | <u>0.0296</u> ± 0.0123 | <u>0.0251</u> ± 0.0172 | 0.0199 ± 0.0179 | <u>0.0296</u> ± 0.0123 | <u>0.0251</u> ± 0.0172 | 0.0199 ± 0.0179 | <u>0.0296</u> ± 0.0123 | <u>0.0251</u> ± 0.0172 | 0.0199 ± 0.0179 |
| SPNet +z | 0.0306 ± 0.0370 | 0.0604 ± 0.0371 | 0.1267 ± 0.0655 | 0.0451 ± 0.0332 | 0.0445 ± 0.0300 | 0.0865 ± 0.0440 | 0.0639 ± 0.0273 | 0.0980 ± 0.0269 | 0.1877 ± 0.0546 | 0.0834 ± 0.0342 | 0.0694 ± 0.0234 | 0.1309 ± 0.0590 |
| Ours | 0.0238 ± 0.0096 | 0.0092 ± 0.0080 | <u>0.0312</u> ± 0.0191 | 0.0235 ± 0.0096 | 0.0091 ± 0.0081 | <u>0.0314</u> ± 0.0190 | 0.0241 ± 0.0094 | 0.0094 ± 0.0080 | <u>0.0312</u> ± 0.0193 | 0.0249 ± 0.0090 | 0.0108 ± 0.0070 | <u>0.0320</u> ± 0.0195 |

Table 4: Experimental results on Flickr(hete) Dataset. The top result is highlighted in bold, and the runner-up is underlined. / means not applicable.

| Methods | $\epsilon_{average}$ | | | | | | $\epsilon_{individual}$ | | | | | |
|----------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| | Within Sample | | | Out-of Sample | | | Within Sample | | | Out-of Sample | | |
| | AME | ASE | ATE | AME | ASE | ATE | IME | ISE | ITE | IME | ISE | ITE |
| TARNET+z | 0.1315 ± 0.0740 | 0.1673 ± 0.0423 | 0.3590 ± 0.0785 | 0.1554 ± 0.1110 | 0.1319 ± 0.0367 | 0.2728 ± 0.1321 | 0.2320 ± 0.0432 | 0.2042 ± 0.0479 | 0.4254 ± 0.0862 | 1.5274 ± 1.6256 | 0.1700 ± 0.0351 | 1.5957 ± 1.5779 |
| CFR+z | 0.1131 ± 0.0476 | 0.1437 ± 0.0081 | 0.2960 ± 0.0182 | 0.0998 ± 0.0458 | 0.1412 ± 0.0085 | 0.2789 ± 0.0309 | 0.1445 ± 0.0407 | 0.1463 ± 0.0083 | 0.3242 ± 0.0224 | 0.5946 ± 0.7379 | 0.1448 ± 0.0087 | 0.7104 ± 0.6761 |
| GEat | 0.3283 ± 0.0426 | 0.4717 ± 0.1336 | 0.1074 ± 0.0255 | 0.3356 ± 0.0279 | 0.4723 ± 0.1312 | 0.0969 ± 0.0099 | 0.3697 ± 0.0396 | 0.5123 ± 0.1231 | 0.2178 ± 0.0214 | 0.7124 ± 0.6463 | 0.5144 ± 0.1202 | 0.5914 ± 0.7073 |
| ND+z | 0.2420 ± 0.0330 | 0.0293 ± 0.0113 | 0.0852 ± 0.0365 | 0.2433 ± 0.0539 | 0.0318 ± 0.0134 | 0.0785 ± 0.0422 | 0.2571 ± 0.0348 | 0.0430 ± 0.0040 | 0.1607 ± 0.0188 | 0.5156 ± 0.2180 | 0.0669 ± 0.0059 | 0.4720 ± 0.2580 |
| NetEst | 0.0530 ± 0.0423 | 0.0452 ± 0.0351 | 0.0723 ± 0.0319 | 0.0466 ± 0.0322 | 0.0565 ± 0.0454 | 0.0818 ± 0.0379 | 0.1145 ± 0.0278 | 0.0607 ± 0.0267 | 0.1660 ± 0.0161 | 0.6855 ± 0.2607 | 0.5353 ± 0.2507 | 0.5625 ± 0.1367 |
| TNet | <u>0.0411</u> ± 0.0238 | 0.0206 ± 0.0073 | 0.0282 ± 0.0297 | <u>0.0417</u> ± 0.0237 | <u>0.0196</u> ± 0.0098 | 0.0268 ± 0.0314 | <u>0.0936</u> ± 0.0170 | <u>0.0338</u> ± 0.0065 | <u>0.1360</u> ± 0.0219 | <u>0.0950</u> ± 0.0157 | <u>0.0361</u> ± 0.0074 | <u>0.1412</u> ± 0.0223 |
| RRNET | 0.0441 ± 0.0196 | 0.0280 ± 0.0153 | <u>0.0367</u> ± 0.0243 | 0.0446 ± 0.0183 | 0.0292 ± 0.0105 | 0.0329 ± 0.0255 | 0.0940 ± 0.0166 | 0.0401 ± 0.0105 | 0.1373 ± 0.0188 | 0.0952 ± 0.0153 | 0.0436 ± 0.0109 | 0.1419 ± 0.0202 |
| SPNet+z | 0.0429 ± 0.0461 | 0.0885 ± 0.0509 | 0.2359 ± 0.0724 | 0.0509 ± 0.0484 | 0.0555 ± 0.0264 | 0.1483 ± 0.0573 | 0.1296 ± 0.0241 | 0.1593 ± 0.0314 | 0.3538 ± 0.0513 | 0.1827 ± 0.1051 | 0.1053 ± 0.0208 | 0.2879 ± 0.0823 |
| Ours | 0.0329 ± 0.0282 | 0.0094 ± 0.0050 | 0.0368 ± 0.0107 | 0.0339 ± 0.0289 | 0.0084 ± 0.0056 | <u>0.0320</u> ± 0.0152 | 0.0925 ± 0.0102 | 0.0284 ± 0.0038 | 0.1362 ± 0.0132 | 0.0942 ± 0.0088 | 0.0310 ± 0.0040 | 0.1412 ± 0.0140 |