

Table 1: Ablation study results on Flickr(homo) dataset. Ours_w/o_CRL& X_N is our method without the causal representation module and only using X as adjustment variables. Ours_w/o_CRL is our method without the causal representation module. Ours_w/o_ELBO&SM is our method setting the coefficient of $\mathcal{L}_{ELBO} + \mathcal{L}_{SM}$ to zero. Ours_w/o_IPM is our method setting the coefficient of \mathcal{L}_{IPW} to zero. Ours (incorrect dim U) is our method setting all dimensions of U^i, U^c, U^n as 16 (ground truth is 3,4,3 respectively).

Methods	$\mathcal{E}_{average}$						$\mathcal{E}_{individual}$					
	Within Sample			Out-of Sample			Within Sample			Out-of Sample		
	AME	ASE	ATE	AME	ASE	ATE	IME	ISE	ITE	IME	ISE	ITE
Ours_w/o_CRL& X_N	0.1960 ± 0.0080	0.5023 ± 0.0000	0.8043 ± 0.0080	0.1806 ± 0.0079	0.4908 ± 0.0001	0.8378 ± 0.0078	0.7170 ± 0.0062	0.8563 ± 0.0048	1.0676 ± 0.0058	0.7468 ± 0.0077	0.8424 ± 0.0047	1.1005 ± 0.0053
Ours_w/o_CRL	0.1937 ± 0.1092	0.4576 ± 0.0485	0.5531 ± 0.0711	0.2083 ± 0.2122	0.4577 ± 0.0498	0.5198 ± 0.0608	0.2135 ± 0.1999	0.4632 ± 0.0535	0.5721 ± 0.0595	0.4279 ± 0.3143	0.4910 ± 0.0929	0.6700 ± 0.1424
Ours_w/o_ELBO&SM	0.0316 ± 0.0172	0.0191 ± 0.0126	0.0257 ± 0.0190	0.0316 ± 0.0172	0.0191 ± 0.0126	0.0258 ± 0.0189	0.0316 ± 0.0172	0.0191 ± 0.0126	0.0257 ± 0.0190	0.0316 ± 0.0172	0.0191 ± 0.0126	0.0258 ± 0.0189
Ours_w/o_IPM	0.0359 ± 0.0202	0.0133 ± 0.0050	0.0598 ± 0.0396	0.0394 ± 0.0282	0.0123 ± 0.0060	0.0602 ± 0.0352	0.0410 ± 0.0228	0.0155 ± 0.0033	0.0643 ± 0.0337	0.0424 ± 0.0242	0.0142 ± 0.0044	0.0632 ± 0.0334
Ours	0.0238 ± 0.0096	0.0092 ± 0.0080	0.0312 ± 0.0193	0.0235 ± 0.0096	0.0091 ± 0.0081	0.0314 ± 0.0190	0.0241 ± 0.0094	0.0094 ± 0.0080	0.0315 ± 0.0193	0.0249 ± 0.0100	0.0108 ± 0.0070	0.0329 ± 0.0195
Ours (incorrect dim U)	0.0247 ± 0.0164	0.0293 ± 0.0153	0.0317 ± 0.0221	0.0254 ± 0.0179	0.0284 ± 0.0153	0.0312 ± 0.0219	0.0253 ± 0.0161	0.0295 ± 0.0153	0.0319 ± 0.0220	0.0256 ± 0.0178	0.0284 ± 0.0153	0.0313 ± 0.0218