We thank the reviewers for the valuable suggestions and appreciate the positive feedback.

To Reviewer #1:

Framing of problem. We agree that our method indeed improves the effective training sample size and reduces the variance of counterfactual estimates. In this paper, we refer confounding bias to the correlation between $\mathbf{X}$ and $\mathbf{T}$ in the training data which causes the distribution shift of the test distribution where $\mathbf{T}$ is randomly assigned. We will further improve the problem framing in new version..

To Reviewer #2:

Correctness of Theorem 1. There might be some misunderstanding here. $P_W(\mathbf{X}, \mathbf{T}) = W_T(\mathbf{X}, \mathbf{T})P(\mathbf{X}, \mathbf{T})$ is defined to be a valid distribution probability density if $\mathbb{E}_{\mathbf{X},\mathbf{T}\sim P(\mathbf{X},\mathbf{T})}[W_T(\mathbf{X}, \mathbf{T})] = 1$. Our sample weights $W_T(\mathbf{X}, \mathbf{T}) = \frac{P(\mathbf{T})}{P(\mathbf{T}|\mathbf{X})}$ shown in line 118 can guarantee this property. And in the Lines 100, 102, our definition of countefactual risk $\mathcal{E}_{cf}$ implies that our testing distribution is $P_{cf}(\mathbf{X}, \mathbf{T}) = P(\mathbf{X})P(\mathbf{T})$, which is also a valid probability density. Therefore, the integral probability metric is defined on valid distributions.

Novelty of our method. We double-checked the literature [6,23], finding no technical content on variational sample weight learning in these works. As far as we know, we are the first to explore the low dimensional latent structure of bundle treatments and variational sample reweighting is proposed to address the challenge of decorrelating X and T.

Is the objective function in supplementary material a updated version? The objective function in the supplementary material is not the updated version of our method. It is a baseline for comparison.

To Reviewer #3:

Comparison to extension of methods in single treatment settings. Many conventional methods of multi-level or binary treatment build prediction model for each treatment and balance the confounder distribution of each treatment group. In the bundle treatment setting, the number of groups can be large and the number of samples in each group can be few. There may even be such cases that when we predict outcome of some treatment $\mathbf{t}$ in test set, there are no training sample in the corresponding group. Therefore, the data insufficiency problem make these methods infeasible. There are some conventional methods that view treatments as input features to make prediction such as BNN in "Learning representations for counterfactual inference". The baselines DNN&$\mathrm{W}_{\mathrm{raw}}$ and DNN&IR can be viewed as extension of these conventional methods incorporating the sample re-weighting and treatment invariant representation learning.

To Reviewer #4:

Details of datasets The matrices A and B are generated from standard gaussian distribution, that is $a_{i,j}, b_{i,j} \sim \mathcal{N}(0, 1)$. We will refine the details and release the code and dataset in the future.

Are Distribution of training and test set the same? Since we randomly shuffled the matches of confounders and treatments, the treatments are randomly assigned in test set. Meanwhile, the confounders and treatments are correlated in training distribution. Therefore, the distribution of training set and test set are not the same.

Are our method limited to additive treatment effect? We check the outcome generation in Recsim dataset is of the form $y = 1 - \frac{1}{1+\sum_i P_i}$, where P_i is determined by the attributes of user and i^{th} document. This does not belong to additive treatment effect. Also, our method is designed without requirement for additive treatment effect model.

Performance when s is not constant. We define $t_j = 1$ if $f_j > e$, otherwise $t_j = 0$. Then the number of treatments is not constant. We set threshold $e = p/10 - 1$. See results in Table 1.

Performance when some information is missing. We conduct experiments when the dimension of latent representation k' in VAE is mismatched with true k. The results in Table 1 shows that our method is robust to this mismatch, even when k' is smaller than true k, demonstrating the effectiveness of our method in the case of missing information.

Whether we know which document is clicked. In our setup, we do not know which document is clicked. The probability of clicking bundle is simultaneously determined by all the documents in bundle and the user.

Fix sample size $n = 10000$, varying dimension of treatments p, non-constant s								
p	$p = 10$		$p = 20$		$p = 30$		$p = 50$	
Methods	Mean	STD	Mean	STD	Mean	STD	Mean	STD
DNN	0.884	0.121	1.126	0.097	1.828	0.130	3.244	0.447
DNN&$\mathrm{W}_{\mathrm{raw}}$	0.778	0.081	1.100	0.078	1.712	**0.117**	3.143	0.288
DNN&W_{AE}	0.775	0.084	1.125	**0.064**	1.618	0.121	3.108	0.271
DNN&$\mathrm{W}_{\mathrm{VSR}}$	**0.657**	**0.068**	**0.975**	0.086	**1.418**	0.125	**2.898**	**0.211**
DNN&IR	0.893	0.107	1.131	0.136	1.819	0.145	3.200	0.273
Fix sample size $n = 10000$, dimension of treatments $p = 10$, varying k'								
k'	$k' = 2$		$k' = 3 = k$		$k' = 4$		$k' = 5$	
Methods	Mean	STD	Mean	STD	Mean	STD	Mean	STD
DNN&$\mathrm{W}_{\mathrm{VSR}}$	0.506	0.051	0.476	0.037	0.479	0.032	0.484	0.050

Table 1: Experiment results for non-constant s and varying k'.