
Towards Balanced Representation Learning for Credit Policy Evaluation

Yiyan Huang
yiyhuang3-c@my.cityu.edu.hk
City University of Hong Kong

Zhiri Yuan
yuanzhiri2012@gmail.com
City University of Hong Kong

Dongdong Wang
wangdongdong9@jd.com
JD Digits

Cheuk Hang Leung
chleung87@cityu.edu.hk
City University of Hong Kong

Qi Wu*
qi.wu@cityu.edu.hk
City University of Hong Kong

Shumin Ma
shuminma@uic.edu.cn
BNU-HKBU United International College

Siyi Wang
swang348-c@my.cityu.edu.hk
City University of Hong Kong

Zhixiang Huang
huangzhixiang@jd.com
JD Digits

Abstract

Credit policy evaluation presents profitable opportunities for E-commerce platforms through improved decision-making. The core of policy evaluation is estimating the causal effects of the policy on the target outcome. However, selection bias presents a key challenge in estimating causal effects from real-world data. Some recent causal inference methods attempt to mitigate selection bias by leveraging covariate balancing in the representation space to obtain the domain-invariant features. However, it is noticeable that balanced representation learning can be accompanied by a failure of domain discrimination, resulting in the loss of domain-related information. This is referred to as the over-balancing issue. In this paper, we introduce a novel objective for representation balancing methods to do policy evaluation. In particular, we construct a doubly robust loss based on the predictions of treatment and outcomes, serving as a prerequisite for covariate balancing to deal with the over-balancing issue. In addition, we investigate how to improve treatment effect estimations by exploiting the unconfoundedness assumption. The extensive experimental results on benchmark datasets and a newly introduced credit dataset show a general outperformance of our method compared with existing methods.

1 INTRODUCTION

The rapid-growing E-commerce has become an indispensable part of the retail industry in the recent decade. With a rising trend of online shopping, many E-commerce platforms have been offering the lending service, which enables approved buyers to utilize their credit line to undergo online payment. When big promotional events such as Black Friday and Double 11 are coming, the E-commerce platform usually raises some consumers' credit limits to encourage their spending. In reality, different people can react very differently to the same credit policy. For example, people with higher consumption demand but with lower credit limits can be sensitive to the credit increase policy. In this scenario, personalized credit policy impact evaluation becomes a vital real-world question for the credit loaner.

The core of policy evaluation is to estimate *causal effects* Heckman and Vytlačil (2005). Taking online consumer finance as an example, if the credit increase policy is the treatment and the default status is the outcome, then our target is to evaluate the direct effect of the credit policy on default status, which is called the *treatment effect*. An important target of policy evaluation problems is to estimate the individual treatment effect (ITE, aka heterogeneous treatment effect) and the average treatment effect (ATE). ITE represents the treatment effect for a specific individual, while ATE is the average of ITE at the population level. Obtaining treatment effects requires the answer to a hypothetical question: What would the outcome (default status) be if one had received an alternative treatment (credit policy)? Such a hypothesized outcome is referred to as the *counterfactual* outcome.

With vast amounts of data collected, researchers have been devoting themselves to effective machine learning methods to estimate their interested quantities. Still, it is difficult for classical machine learning methods to deal with problems like policy evaluation. This is because though the advanced models have conceivable predictive power,

Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, Valencia, Spain. PMLR: Volume 206. Copyright 2023 by the author(s).

* Corresponding author.

they cannot handle *selection bias*. For example, the actual cause-and-effect relation is that a higher credit line leads to a higher default probability, but a typical regression model can infer a relation that the higher the credit line, the lower the default probability Huang et al. (2020). Such a phenomenon exists because the default probability is affected by some factors that are also used for credit policy assignments. These factors are named confounders/covariates, and they can incur selection bias due to the non-random policy assignment, which presents a major challenge in estimating treatment effects.

To eliminate selection bias, the standard way is to conduct randomized controlled trials (RCTs). However, RCTs are infeasible for E-commerce credit policy evaluation since RCTs are very expensive and time-consuming. More realistically, researchers attempt to use models that directly estimate treatment effects from observational data, such as the outcome model. The outcome model infers (covariates, treatment)-outcome relation, but it is subject to the covariate shift problem incurred by selection bias Shalit et al. (2017). Another mainstream method considers balancing the distribution of covariates by weighting them Rosenbaum and Rubin (1983); Lunceford and Davidian (2004); Li et al. (2018), and the weight usually involves the treatment-outcome relation, which is known as the *propensity score*.

Recently, the neural network models have boosted the representation balancing methods Johansson et al. (2016); Shalit et al. (2017). These methods aim to achieve domain-invariant representations by enforcing the distribution of covariates between the treated and controlled groups to be balanced in the representation space. However, on the one hand, previous causal inference methods are mainly based on *unconfoundedness* assumption, whereas this assumption has not been fully exploited in practice; on the other hand, if a neural net model considers merely the *domain invariance* instead of *domain discrimination*, the model might suffer an over-balancing issue and lose the domain-discriminative (domain-related) information that contributes to the treatment and outcome modeling Guo et al. (2020); Zhang et al. (2020); Assaad et al. (2021); Huang et al. (2022a). To explicate this phenomenon more intuitively, we give a toy example below with illustrations in Figure 1.

Toy example. Let X denote the covariate (monthly consumption), Y denote the outcome (default status), $T = 1$ denote the treatment (increasing credit limit), and $T = 0$ denote the control (no change to credit limit). Suppose that the credit policy is assigned according to the consumption in the last month. The outcome is generated by $Y = 1$ (default) if $P(T = 1|X = x) > 0.5$, otherwise $Y = 0$ (pay off). The left graph in Figure 1 indicates that the higher monthly consumption, the higher probability of one getting

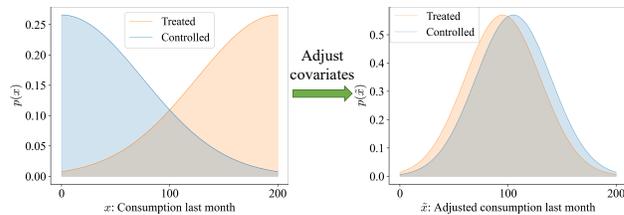


Figure 1: Toy example for illustrating the tradeoff between domain invariance and domain discrimination.

default. In contrast, as shown in the right panel of Figure 1, the distribution of \tilde{X} (\tilde{X} is the adjusted covariates based on X) is invariant between treatment and control domains. In this case, however, precisely predicting T (treatment or control) becomes more difficult. This is because the distribution of \tilde{X} is so balanced that we can hardly distinguish between treated and controlled units. Consequently, it is hard to predict Y accurately using the over-balancing features \tilde{X} . Therefore, although domain invariance can alleviate selection bias, improperly balanced features can be indistinguishable, thereby failing to preserve the domain-related information that is helpful for treatment and outcome predictions.

Contribution. This paper proposes a novel balanced representation learning method, DRRB-OB, to evaluate the effect of credit increase policy on consumption from a newly introduced observational dataset, Credit, at the individual level (ITE) and population level (ATE). In particular, aforementioned issues motivate us to introduce a framework involving doubly robust representation balancing and outcome balancing (DRRB-OB). The main contributions are fourfold: (1) We propose a doubly robust loss as a prerequisite for domain invariance, which prevents the learned representation from being over-balanced. This idea can also be extended to other machine learning areas, such as unsupervised domain adaptation. (2) We fully exploit the advantage of the unconfoundedness assumption by proposing an outcome balancing loss to improve treatment effect estimations. (3) Our method is applicable to the online credit policy evaluation at either the individual level or population level. (4) We create a new Credit dataset that mimics the distribution of real-world online consumer data. This Credit dataset will be very useful for future credit limit management and policy evaluation studies.

2 BACKGROUND

2.1 Potential Outcome Framework

Consider N i.i.d. samples $\{(\mathbf{X}_i, T_i, Y_i)\}_{i=1}^N$. Let the random variable $\mathbf{X}_i \in \mathcal{X} \subset \mathbb{R}^s$ be s -dimensional confounders/covariates, $T_i \in \{0, 1\}$ be the binary treatment, and $Y_i^1, Y_i^0 \in \mathcal{Y} \subset \mathbb{R}$ be the potential outcomes for treat-

ment $T_i = 1$ and $T_i = 0$, respectively. Under the potential outcome (PO) framework (aka, Rubin Causal Model) Rubin (2005), the treatment assignment is according to the propensity score $P(T_i = 1 | \mathbf{X}_i)$, and the factual outcome Y_i is determined by $Y_i = T_i Y_i^1 + (1 - T_i) Y_i^0$. Let (\mathbf{x}_i, t_i, y_i) be the i^{th} observed realization of (\mathbf{X}_i, T_i, Y_i) . Our target is to estimate ITE and ATE, denoted by $\tau_{ITE}(\mathbf{x})$ and τ_{ATE} , respectively:

$$\tau_{ITE}(\mathbf{x}) = \mathbb{E}[Y^1 - Y^0 | \mathbf{X} = \mathbf{x}], \quad (1)$$

$$\tau_{ATE} = \mathbb{E}[Y^1 - Y^0] = \mathbb{E}[\mathbb{E}[Y^1 - Y^0 | \mathbf{X}]]. \quad (2)$$

Identifying ITE and ATE requires the following assumptions under the PO framework:

Assumption 1 (Consistency). *If the treatment is t , then the observed outcome equals the potential outcome under the treatment t : When $T = t$, $Y = Y^t \forall t \in \{0, 1\}$.*

Assumption 2 (Overlap). *The propensity score is bounded away from 0 to 1: $0 < P(T = 1 | \mathbf{X} = \mathbf{x}) < 1$ and $0 < P(T = 0 | \mathbf{X} = \mathbf{x}) < 1 \forall \mathbf{x} \in \mathbb{R}^s$.*

Assumption 3 (Unconfoundedness). *The potential outcomes are independent of treatment assignment given covariates: $Y^t \perp\!\!\!\perp T | \mathbf{X}, \forall t \in \{0, 1\}$.*

After obtaining the estimates $\hat{\tau}_{ITE}(\mathbf{x})$ and $\hat{\tau}_{ATE}$, we can measure the model performance on policy evaluation by the following metrics ϵ_{PEHE} (PEHE stands for Precision in Estimation of Heterogeneous Effects) and ϵ_{ATE} :

$$\epsilon_{PEHE} = \int_{\mathcal{X}} (\hat{\tau}_{ITE}(\mathbf{x}) - \tau_{ITE}(\mathbf{x}))^2 p(\mathbf{x}) d\mathbf{x}, \quad (3)$$

$$\begin{aligned} \epsilon_{ATE} &= \left| \int_{\mathcal{X}} (\hat{\tau}_{ITE}(\mathbf{x}) - \tau_{ITE}(\mathbf{x})) p(\mathbf{x}) d\mathbf{x} \right| \\ &= |\hat{\tau}_{ATE} - \tau_{ATE}|. \end{aligned} \quad (4)$$

2.2 Related Works

Doubly robust estimation. The doubly robust (DR) estimator has been thoroughly studied for ATE estimation in many pieces of literature. For example, the debiased machine learning (DML) estimator (Chernozhukov et al., 2018) is doubly robust for ATE estimation. The DML estimator is consistent if either the outcome model or propensity score model is correctly specified with a specific convergence rate. Additionally, the DR-learner (Kennedy, 2020) is designed with doubly robust pseudo-outcomes for ITE estimation. The DR-learner requires two-stage estimations: First, the DR-learner fits the outcome model and the propensity score model to construct the so-called doubly robust pseudo-outcomes. Second, it regresses the difference of constructed outcomes on covariates and predicts the ITE directly. Traditional machine learning methods, e.g., logistic regression and tree models, can also be used to evaluate credit policy, and they have been extended under the debiased machine learning framework (Oprescu et al.,

2019; Liu et al., 2021) Other studies in terms of doubly robust estimators or the debiased machine learning method include Farrell (2015); Yang et al. (2020); Huang et al. (2020); Knaus (2021); Jung et al. (2021); Huang et al. (2022b) and references therein.

Representation learning. Our work has a strong connection with representation learning methods for causal inference. These methods share very similar basic neural net structures (Shalit et al., 2017; Du et al., 2021; Assaad et al., 2021). The basic architecture of these methods consists of a representation network $\Phi : \mathcal{X} \rightarrow \mathcal{R}$ and two-head outcome networks $f_1 : \mathcal{X} \times \{1\} \rightarrow \mathcal{Y}$ and $f_0 : \mathcal{X} \times \{0\} \rightarrow \mathcal{Y}$. Some methods (Shi et al., 2019) also consider an additional propensity head network $\pi : \mathcal{R} \rightarrow \mathbb{R}$. For causal representation learning methods, they usually achieve domain-invariant representations by minimizing the distribution distance between treated and controlled groups in the representation space. The distance is measured by the integral probability metric (IPM), denoted by $IPM_G(\cdot, \cdot)$, where G is a function family consisting of functions $g : \mathcal{R} \rightarrow \mathbb{R}$. If we assume that G is the functional space of a family of 1-Lipschitz functions, then IPM_G becomes 1-Wasserstein distance, denoted by $Wass(\cdot, \cdot)$. This idea is inspired from domain adaptation (Ben-David et al., 2006, 2010) The various representation learning methods (Louizos et al., 2017; Yao et al., 2018; Johansson et al., 2016; Yoon et al., 2018; Zhang et al., 2020; Assaad et al., 2021; Du et al., 2021) can be found in Section 4 and Appendix.

3 REPRESENTATION BALANCING

In this section, we first state the theoretical generalization bounds for balanced representation learning models in Section 3.1. The over-balancing problem is then demonstrated in Section 3.2. Section 3.3 next introduces the idea of involving the domain discrimination task in representation learning models. Finally, Section 3.4 discusses how to resolve the over-balancing issue and exploit the benefit of the unconfoundedness assumption, under the proposed DRRB framework.

3.1 Generalization Bounds

Due to the existence of selection bias, the covariate distribution shift problem is widespread in observational E-commerce data. For example, a credit increase policy is more likely to assign to individuals who have high historical consumption and low default risks. This indicates that the covariate distributions of the treated and controlled groups can be considerably imbalanced. Therefore, we propose to use representation balancing to tackle this issue. Below we present the theoretical generalization bounds in representation balancing methods (Shalit et al., 2017).

Theorem 1. *Let Ψ be the inverse of Φ , $p_{\Phi}^{T=1}$*

and $p_{\Phi}^{T=0}$ be the densities of the treated and controlled covariates in the representation space, $h : \mathcal{R} \times \{0, 1\} \rightarrow \mathcal{Y}$ such that $h(\Phi(\mathbf{x}), t)$ estimates y^t , and G be the family of 1-Lipschitz functions. Assume that there exists a constant $B_{\Phi} \geq 0$ such that $g_{\Phi, h}(\mathbf{r}, t) := \frac{1}{B_{\Phi}} \cdot \ell_{h, \Phi}(\Psi(\mathbf{r}), t) \in G$ for any $t \in \{0, 1\}$, where $\ell_{h, \Phi}(\mathbf{x}, t) = \int_{\mathcal{Y}} (y^t, h(\Phi(\mathbf{x}), t))^2 p(y^t | \mathbf{x}) dy^t$. Defining $\sigma_y^2 = \min\{\sigma_{y^t}^2(p(\mathbf{x}, t)), \sigma_{y^t}^2(p(\mathbf{x}, 1 - t))\} \forall t \in \{0, 1\}$, where $\sigma_{y^t}^2(p(\mathbf{x}, t)) = \int_{\mathcal{X} \times \{0, 1\} \times \mathcal{Y}} (y^t - \tau^t(\mathbf{x}))^2 p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt$, we then have

$$\begin{aligned} \epsilon_{PEHE}(h, \Phi) &\leq 2(\epsilon_F^{T=1}(h, \Phi) + \epsilon_F^{T=0}(h, \Phi)) \\ &\quad + B_{\Phi} \cdot IPM_G(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}) - 2\sigma_y^2. \end{aligned} \quad (5)$$

Here, $\epsilon_F^{T=t}(h, \Phi) = \int_{\mathcal{X}} \ell_{h, \Phi}(\mathbf{x}, t) p^{T=t}(\mathbf{x}) d\mathbf{x}$ is the estimation error over factual outcomes for $t \in \{0, 1\}$. $IPM_G(p_{\Phi}^{T=1}, p_{\Phi}^{T=0})$ measures the distributional discrepancy between the treated domain $P(\Phi(\mathbf{X}) | T = 1)$ and the controlled domain $P(\Phi(\mathbf{X}) | T = 0)$. We defer the proof of Theorem 1 to Appendix.

3.2 The Conundrum in Representation Balancing

The generalization bound for ϵ_{PEHE} in (5) indicates that we can handle selection bias by (i) achieving a low prediction error over factual outcomes and (ii) obtaining domain-invariant representations by minimizing the distribution discrepancy between the treated and controlled groups in the representation space.

To achieve the goal (i), we define $f(\cdot, t_i) = t_i f_1(\cdot) + (1 - t_i) f_0(\cdot)$ as the estimate of factual outcome, where f_1 and f_0 estimate the potential outcome y^1 and y^0 , respectively. We aim to minimize the loss over factual outcomes:

$$\mathcal{L}_y(\mathbf{x}_i, t_i, y_i; \Phi, f) = [f(\Phi(\mathbf{x}_i), t_i) - y_i]^2. \quad (6)$$

The goal (ii) aims to find a representation function Φ such that Φ minimizes the distance between the treated domain and the controlled domain in the representation space:

$$IPM_G(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}). \quad (7)$$

If IPM_G is chosen as 1-Wasserstein distance, then the final objective of the representation balancing models is

$$\min_{\Phi, f} \frac{1}{N} \sum_{i=1}^N \mathcal{L}_y(\mathbf{x}_i, t_i, y_i; \Phi, f) + \alpha_1 \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}). \quad (8)$$

The most representative domain-invariant representation learning model is CFR-Wass (Shalit et al., 2017) (see Figure 2(a)). However, representation learning models that merely consider domain invariance may suffer an over-balancing issue since the objective (8) imposes no restrictions on the domain invariance task. As a result, Φ

would keep updating for the goal of domain invariance until the optimization converges. In this scenario, Φ is prone to being over-balanced and hence may lose the domain-discriminative information useful for outcome modeling.

3.3 Combining with Domain Discriminator

The toy example in Figure 1 demonstrates the inadequacy of domain invariance to preserve domain-related information to prevent representation balancing models from the over-balancing problem. In this section, we will introduce the domain discrimination task for the representation learning methods.

Direct objective. An intuitive and direct approach to incorporating domain-discriminative information is to force the model to learn domain-discriminative representations. This allows the treated units to be easily distinguished from the controlled ones in the representation space. Specifically, we can minimize the treatment prediction error measured by the cross-entropy loss $\mathcal{L}_t(\mathbf{x}_i, t_i; \Phi, \pi)$ by jointly training the domain discriminator (treatment classifier) π and the representation function Φ such that

$$\begin{aligned} \mathcal{L}_t(\mathbf{x}_i, t_i; \Phi, \pi) &= -t_i \log \pi(\Phi(\mathbf{x}_i)) \\ &\quad - (1 - t_i) \log(1 - \pi(\Phi(\mathbf{x}_i))). \end{aligned} \quad (9)$$

Together with the outcome loss in (6), a direct objective of domain-discriminative models is

$$\min_{\Phi, f, \pi} \frac{1}{N} \sum_{i=1}^N [\mathcal{L}_y(\mathbf{x}_i, t_i, y_i; \Phi, f) + \alpha_2 \mathcal{L}_t(\mathbf{x}_i, t_i; \Phi, \pi)]. \quad (10)$$

One of the representative domain-discriminative representation learning models is called Dragonnet (Shi et al., 2019) (see Figure 2(b)). Such a domain-discriminative model can preserve domain-related patterns, but the objective makes it hard to deal with the selection bias problem due to lacking the domain-invariant task.

Multi-task objective. As discussed above, merely desiring domain invariance might cause the model to fall into an over-balancing issue, while considering domain discrimination alone is insufficient to handle selection bias. Naturally, we can combine both domain invariance (8) and domain discrimination (10) with a multi-task learning objective:

$$\begin{aligned} \min_{\Phi, f, \pi} \frac{1}{N} \sum_{i=1}^N \mathcal{L}_y(\mathbf{x}_i, t_i, y_i; \Phi, f) + \alpha_1 \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}) \\ + \alpha_2 \frac{1}{N} \sum_{i=1}^N \mathcal{L}_t(\mathbf{x}_i, t_i; \Phi, \pi). \end{aligned} \quad (11)$$

Heuristically, the multi-task objective (11) forces Φ to preserve domain-discriminative information using a domain

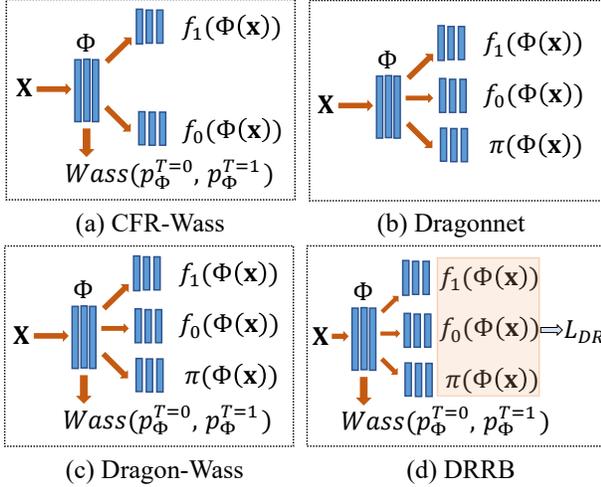


Figure 2: Illustrations of CFR-Wass, Dragonnet, Dragon-Wass, and the proposed DRRB methods.

discriminator, protecting representation balancing models from the over-balancing problem. As a combination of CFR-Wass and Dragonnet, the model with the objective (11) can be referred to as Dragon-Wass (see Figure 2(c)).

The multi-task objective (11) is expected to compensate for the shortfalls of CFR-Wass and Dragonnet, but there still remain some challenges confronted by Dragon-Wass. First, the fact that domain invariance and domain discrimination counteract each other makes it difficult to strike a balance between the two subtasks, as the optimal trade-off is usually unknown in practice. Second, the objective (11) just combines the three sub-objectives together, without explicating the interaction between them. Third, like other existing causal representation learning methods, Dragon-Wass does not make use of the unconfoundedness assumption.

3.4 Doubly Robust Representation Balancing

The dilemmas encountered by the above representation learning models motivate us to explore a new representation balancing framework, the **doubly robust representation balancing (DRRB)** framework. Specifically, we construct the doubly robust loss $\mathcal{L}_{DR}(\mathbf{x}_i, t_i, y_i; \Phi, f, \pi)$ to avoid the over-balancing issue, and establish the outcome balancing loss $\mathcal{L}_{OB}(\mathbf{x}, t; \Phi, f_0, f_1)$ to adapt the model to the unconfoundedness assumption stated in Assumption 3. We first introduce the doubly robust loss and explain how it avoids the over-balancing issue.

Doubly robust loss. Recall that domain-related information is associated with the treatment and outcome modeling. Therefore, it can be concluded that *if either treatment or outcome is accurately predicted, domain-related information is considered well-preserved (i.e., no over-balancing issue)*. Inspired by this doubly robust prop-

erty, we construct the pseudo-outcomes $y_1^{DR}(t_i, \cdot)$ and $y_0^{DR}(t_i, \cdot)$ such that

$$\begin{aligned} y_1^{DR}(t_i, \cdot) &= f_1(\cdot) + \frac{t_i}{\pi(\cdot)}(y_i - f_1(\cdot)); \\ y_0^{DR}(t_i, \cdot) &= f_0(\cdot) + \frac{1-t_i}{1-\pi(\cdot)}(y_i - f_0(\cdot)). \end{aligned} \quad (12)$$

Based on (12), we have the doubly robust loss

$$\begin{aligned} \mathcal{L}_{DR}(\mathbf{x}_i, t_i, y_i; \Phi, f, \pi) &= [t_i y_1^{DR}(t_i, \mathbf{x}_i) + (1-t_i) y_0^{DR}(t_i, \mathbf{x}_i) - y_i]^2 \\ &= t_i (f_1(\Phi(\mathbf{x}_i)) - y_i)^2 \left(1 - \frac{t_i}{\pi(\Phi(\mathbf{x}_i))}\right)^2 \\ &\quad + (1-t_i) (f_0(\Phi(\mathbf{x}_i)) - y_i)^2 \left(1 - \frac{1-t_i}{1-\pi(\Phi(\mathbf{x}_i))}\right)^2. \end{aligned} \quad (13)$$

The doubly robust loss gains the following two insights into learning domain-invariant representations. First: Take $t_i = 1$ as an example (we have similar insights for $t_i = 0$). If either $f_1(\Phi(\mathbf{x}_i))$ correctly predicts the outcome or $\pi(\Phi(\mathbf{x}_i))$ accurately estimates the propensity score, we believe that Φ does not lose the domain-related information useful for treatment and outcome modeling, i.e., no over-balancing issue. In this case, the small value of $(1 - \frac{t_i}{\pi(\Phi(\mathbf{x}_i))})^2$ or $(f_1(\Phi(\mathbf{x}_i)) - y_i)^2$ will result in the small sample mean of $\mathcal{L}_{DR}(\mathbf{x}_i, t_i, y_i; \Phi, f, \pi)$, and the representation encoder Φ is allowed to update for the domain invariance task, i.e., minimizing $Wass(p_{\Phi}^{T=1}, p_{\Phi}^{T=0})$. Second: If both $f_1(\Phi(\mathbf{x}_i))$ and $\pi(\Phi(\mathbf{x}_i))$ are not correctly specified, then it indicates that domain-related information might be unintentionally removed in the representation space. That is, the learned representations suffer an over-balancing issue. In this case, the sample mean of $\mathcal{L}_{DR}(\mathbf{x}_i, t_i, y_i; \Phi, f, \pi)$ is large, and the representation encoder Φ should not be updated for the domain invariance task anymore.

Note that quantities of a similar form to the pseudo-outcomes in (12) and (13) have also appeared in the doubly robust estimator Chernozhukov et al. (2018) and DR-learner Kennedy (2020). However, the doubly estimator is a plug-in ATE estimator, and the DR-learner regresses the difference of pseudo-outcomes to estimate ITE, so they are different from how we utilize such pseudo-outcomes. We take advantage of the doubly robust property of these pseudo-outcomes to resolve the over-balancing issue. To be specific, $\mathcal{L}_{DR}(\mathbf{x}_i, t_i, y_i; \Phi, f, \pi)$ will be small if **either** the outcome **or** the treatment is well predicted, but not necessarily both. Therefore, the proposed doubly robust loss can be used as an indicator of the over-balancing problem, which brings a new insight into the representation balancing method.

Outcome balancing loss. Now we are ready to demonstrate how to utilize the unconfoundedness assumption for representation balancing models. Previous causal inference

methods are based on the PO framework, where the unconfoundedness assumption (Assumption 3) is necessary, i.e., $Y^t \perp\!\!\!\perp T \mid \mathbf{X}, \forall t \in \{0, 1\}$. According to the unconfoundedness assumption, we have the following property:

$$P(Y^j \mid \mathbf{X}, T) = P(Y^j \mid \mathbf{X}) \quad \forall j \in \{0, 1\}. \quad (14)$$

The unconfoundedness indicates that the potential outcomes are independent of the treatment assignment given specific covariates. Consequently, the distribution of potential outcome Y^j with $\forall j \in \{0, 1\}$ in treated units should be identical to that in controlled units given $\mathbf{X} \sim \mathcal{X}$. In our model setup, $f_j \circ \Phi(\mathbf{X})$ is the approximate of Y^j . Therefore, we encourage representation balancing models to balance the distributions of potential outcomes $f_j \circ \Phi(\mathbf{X})$ for $j \in \{0, 1\}$ between treated and controlled groups based on the property given in (14). To be specific, we propose the outcome balancing loss $\mathcal{L}_{OB}(\mathbf{x}, t; \Phi, f_0, f_1)$ such that

$$\begin{aligned} \mathcal{L}_{OB}(\mathbf{x}, t; \Phi, f_0, f_1) &= \text{Wass}(p_{f_1}^{T=1}, p_{f_1}^{T=0}) \\ &+ \text{Wass}(p_{f_0}^{T=1}, p_{f_0}^{T=0}). \end{aligned} \quad (15)$$

The final objective. Combining the doubly robust loss (13) and the outcome balancing loss (15), we have the final model **DRRB-OB**. The objective of DRRB-OB involves three separated steps:

$$\text{Step 1: } \min_{\Phi, f_0, f_1} \frac{1}{N} \sum_{i=1}^N \mathcal{L}_y(\mathbf{x}_i, t_i, y_i; \Phi, f) \quad (16)$$

$$+ \lambda_1 \mathcal{L}_{OB}(\mathbf{x}, t; \Phi, f_0, f_1); \quad (17)$$

$$\text{Step 2: } \min_{\pi} \frac{1}{N} \sum_{i=1}^N \mathcal{L}_t(\mathbf{x}_i, t_i; \Phi, \pi); \quad (18)$$

$$\text{if } \mathcal{L}_{DR}(\mathbf{x}_i, t_i, y_i; \Phi, f, \pi) < \lambda_2: \quad (19)$$

$$\text{Step 3: } \min_{\Phi} \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}). \quad (20)$$

Within each iteration, equation (16) accommodates the model to outcome predictions, which corresponds to $\epsilon_F^{T=t}(h, \Phi)$ in equation (5) of Theorem 1; equation (17) adapts the model to the unconfoundedness assumption, which corresponds to $\mathcal{L}_{OB}(\mathbf{x}, t; \Phi, f_0, f_1)$ in equation (15). Further, the domain discriminator is trained in (18) to estimate the propensity score using $\pi(\Phi(\mathbf{x}_i))$. Note that unlike (10) or (11) where π and Φ are updated simultaneously, equation (18) is primarily designed to use π to judge whether $\Phi(\mathbf{x}_i)$ loses domain-discriminative information. In (19), the DRRB method cleverly takes advantage of the doubly robust property to identify whether $\Phi(\mathbf{x}_i)$ suffers an over-balancing issue. In particular, given a threshold λ_2 , if (19) is satisfied, domain-related information is believed to be well-preserved without an over-balancing problem. In this scenario, Φ is allowed to be further updated for the domain invariance task (20), which corresponds to $\text{IPM}_G(p_{\Phi}^{T=1}, p_{\Phi}^{T=0})$ in equation (5) of Theorem 1. Here, λ_1 and λ_2 are hyperparameters, where the

threshold λ_2 can be either fixed or adaptive. Note that the discrepancy measure can be changed to other divergence metrics other than 1-Wasserstein distance. It is also an important step for future studies to solve the time-consuming problem of training our model with multi-task objectives.

4 EXPERIMENTS

This section gives comprehensive experimental results of our method and baseline models on causal benchmark datasets and an E-commerce retail credit dataset. All the experiments are run on Dell 7920 with 1x 16-core Intel Xeon Gold 6250 3.90GHz CPU and 3x NVIDIA Quadro RTX 6000 GPU. In the subsequent investigation, we mainly go deep into two questions: **(1)** Does the proposed method outperform other baseline methods in terms of treatment effect estimation? **(2)** Do the proposed components, i.e., the doubly robust loss in (19) and the outcome balancing loss in (17), contribute to improving treatment effect estimations?

4.1 Experiments on Benchmark Datasets

In reality, the ground truth of treatment effects is unavailable because each individual can only receive treatment 0 or treatment 1. That is, y^1 and y^0 are not available simultaneously. Previous works adopt semi-synthetic datasets to evaluate the performances of causal inference methods, of which **IHDP** and **Twins** are the most frequently used benchmark datasets.

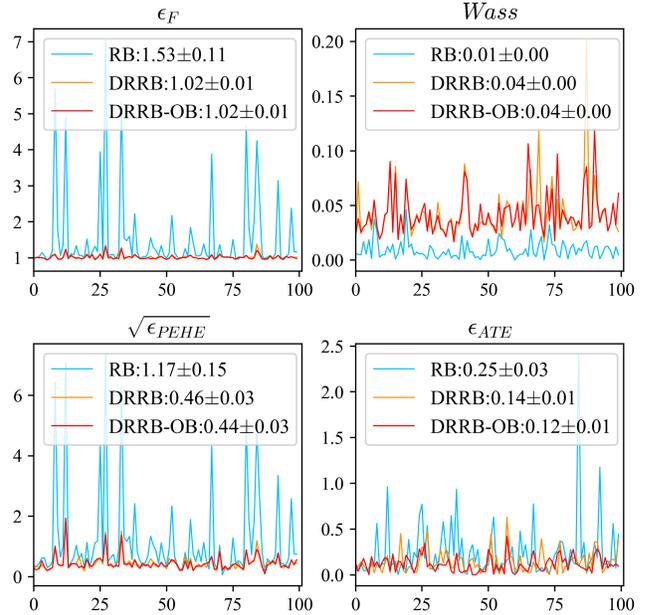
Benchmark datasets. **IHDP:** The IHDP dataset is introduced by Hill (2011). The dataset consists of 747 samples with 25-dimensional covariates collected from a real-world randomized experiment. The selection bias is created by removing a part of treated samples, and the target is to study the effect of special visits (treatment) on cognitive scores (outcome). The potential outcomes are generated using the NPCI package Dorie (2021). We use the same 1000 datasets as used in Shalit et al. (2017), with each dataset split by the ratio of 63%/27%/10% as training/validation/test sets. **Twins:** The Twins dataset contains 11400 pairs of twins weighing less than 2 kg in the USA between 1989 and 1991 Almond et al. (2005). There are 30 attributes associated with parents, pregnancy, and birth as the covariates. The treatment $T = 1$ and $T = 0$ indicate the heavier twin and the lighter twin, respectively. The outcome is 1-year mortality, and the average mortality rate is 17.7% for the lighter twin and 16.1% for the heavier twin. The selection bias is created by choosing one of the two twins as the factual observation Yoon et al. (2018): $T_i \mid \mathbf{X}_i \sim \text{Bernoulli}((1 + \exp(\mathbf{w}^T \mathbf{X}_i + n))^{-1})$, where $\mathbf{w} \sim \mathcal{U}((-0.01, 0.01)^{30 \times 1})$ and $n \sim \mathcal{N}(0, 0.01)$. We repeat the process to generate 100 datasets, with each dataset split by the ratio of 56%/24%/20% as training/validation/test sets.

Table 1: Performance comparisons with mean \pm standard error on benchmark datasets **IHDP** and **Twins**. $\sqrt{\epsilon_{PEHE}}$: Lower is better. ϵ_{ATE} : Lower is better. AUC: Higher is better. Bold indicates the best results across different models.

Method	IHDP				Twins			
	Within-sample		Out-of-sample		Within-sample		Out-of-sample	
	$\sqrt{\epsilon_{PEHE}}$	ϵ_{ATE}	$\sqrt{\epsilon_{PEHE}}$	ϵ_{ATE}	AUC	ϵ_{ATE}	AUC	ϵ_{ATE}
OLS/LR ₁	5.8 \pm .3	.73 \pm .04	5.8 \pm .3	.94 \pm .06	.660 \pm .005	.004 \pm .003	.500 \pm .028	.007 \pm .006
OLS/LR ₂	2.4 \pm .1	.14 \pm .01	2.5 \pm .1	.31 \pm .02	.660 \pm .004	.004 \pm .003	.500 \pm .016	.007 \pm .006
k-NN	2.1 \pm .1	.14 \pm .01	4.1 \pm .2	.79 \pm .05	.609 \pm .010	.003 \pm .002	.492 \pm .012	.005 \pm .004
BART Chipman et al. (2010)	2.1 \pm .1	.23 \pm .01	2.3 \pm .1	.34 \pm .02	.506 \pm .014	.121 \pm .024	.500 \pm .011	.127 \pm .024
CEVAE Louizos et al. (2017)	2.7 \pm .1	.34 \pm .01	2.6 \pm .1	.46 \pm .02	.845 \pm .003	.022 \pm .002	.841 \pm .004	.032 \pm .003
SITE Yao et al. (2018)	.69 \pm .0	.22 \pm .01	.75 \pm .0	.24 \pm .01	.862 \pm .002	.016 \pm .001	.853 \pm .006	.020 \pm .002
BLR Johansson et al. (2016)	5.8 \pm .3	.72 \pm .04	5.8 \pm .3	.93 \pm .05	.611 \pm .009	.006 \pm .004	.510 \pm .018	.033 \pm .009
BNN Johansson et al. (2016)	2.2 \pm .1	.37 \pm .03	2.1 \pm .1	.42 \pm .03	.690 \pm .008	.006 \pm .003	.676 \pm .008	.020 \pm .007
TARNet Shalit et al. (2017)	.88 \pm .0	.26 \pm .01	.95 \pm .0	.28 \pm .01	.849 \pm .002	.011 \pm .002	.840 \pm .006	.015 \pm .002
CFR-Wass Shalit et al. (2017)	.71 \pm .0	.25 \pm .01	.76 \pm .0	.27 \pm .01	.850 \pm .002	.011 \pm .002	.842 \pm .005	.028 \pm .003
Dragonnet Shi et al. (2019)	.89 \pm .0	.15 \pm .01	.93 \pm .0	.18 \pm .01	.880 \pm .000	.005 \pm .005	.874 \pm .001	.008 \pm .005
ABCEI Du et al. (2021)	.96 \pm .0	.18 \pm .01	1.1 \pm .0	.20 \pm .01	.871 \pm .001	.003 \pm .001	.863 \pm .001	.005 \pm .001
Dragon-Wass	.57 \pm .0	.13 \pm .01	.61 \pm .0	.15 \pm .01	.877 \pm .000	.004 \pm .001	.874 \pm .001	.007 \pm .001
DRRB-OB (Ours)	.46 \pm .0	.12 \pm .01	.49 \pm .0	.13 \pm .01	.880 \pm .000	.004 \pm .000	.875 \pm .001	.008 \pm .001

Implementation Details We adopt the absolute error in ATE, $\epsilon_{ATE} = \left| \frac{1}{N} \sum_{i=1}^N ((y_i^1 - y_i^0) - (\hat{y}_i^1 - \hat{y}_i^0)) \right|$, to evaluate the estimation error in ATE. As for the estimation error in ITE, we use Precision in Estimation of Heterogeneous $\epsilon_{PEHE} = \frac{1}{N} \sum_{i=1}^N ((y_i^1 - y_i^0) - (\hat{y}_i^1 - \hat{y}_i^0))^2$ Shalit et al. (2017) for IHDP datasets, and Area Under ROC Curve (AUC) Louizos et al. (2017) for Twins datasets. To analyze the source of gain, we let $\epsilon_F = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$ denote the error in factual outcome estimation, and $Wass$ denote the empirical approximation of $Wass(p_{\Phi}^{T=1}, p_{\Phi}^{T=0})$. The model structure and hyperparameters are set as stated in the implementation details of Appendix.

Ablation study. To investigate the source of gain of the proposed components, we compare DRRB-OB with DRRB (with (17) removed) and RB (with (17) and (19) removed) for various metrics on 1-100 IHDP datasets in Figure 3. The top of each subplot presents the mean \pm standard error of each metric averaged over 100 runs. In Figure 3, we gain an important insight that RB achieves more balanced representations (smaller $Wass$) than DRRB, but such balancing can be harmful to factual outcome estimation (ϵ_F). In contrast, DRRB that involves the doubly robust loss in (17) yields a general outperformance in most cases for RB, reducing ϵ_F by $|1.02/1.53 - 1| = 33.3\%$. Consequently, DRRB attains substantially better ITE estimation ($\sqrt{\epsilon_{PEHE}}$) and ATE estimation (ϵ_{ATE}), with errors reduced by $|0.46/1.17 - 1| = 60.7\%$ for $\sqrt{\epsilon_{PEHE}}$ and $|0.12/0.25 - 1| = 44.0\%$ for ϵ_{ATE} . We also note that the outcome balancing loss in (17) does not affect factual outcome estimation (ϵ_F) or domain invariance ($Wass$), but it improves treatment effect estimation by adapting the model to the unconfoundedness assumption. Specifically, DRRB-OB achieves better ITE and ATE estimations than DRRB, with $|0.44/0.46 - 1| = 4.3\%$ smaller for $\sqrt{\epsilon_{PEHE}}$ and $|0.12/0.14 - 1| = 14.3\%$ smaller for ϵ_{ATE} .


 Figure 3: Plots of model performances on test set for ϵ_F , $Wass$, $\sqrt{\epsilon_{PEHE}}$, and ϵ_{ATE} . Each graph plots the metric for 1-100 IHDP datasets. Mean \pm std of each metric averaged across 100 runs are reported on the top.

Comparison with baselines. We further compare baseline models with DRRB-OB on 1000 IHDP and 100 Twins datasets, and report the model performances (mean \pm standard error) in Table 1. We bold the method with the best performance and defer more detailed comparisons to Appendix. From Table 1, we have the following observations: 1. DRRB-OB achieves superior performances among all the models on IHDP datasets. Notably, DRRB-OB attains significantly smaller treatment effect errors compared to the second-best result, and the error reduction is especially pronounced for $\sqrt{\epsilon_{PEHE}}$, reaching $|0.46/0.57 - 1| =$

Table 2: The comparisons of the mean and standard deviation for the outcome and covariates between the generated data and the real data in treated and controlled groups.

Variables	Mean				Standard Deviation			
	Controlled		Treated		Controlled		Treated	
	Fake	Real	Fake	Real	Fake	Real	Fake	Real
Age	31.02	30.99	30.54	30.00	7.09	6.68	7.41	6.92
Used days	950.10	921.28	874.61	875.62	324.99	310.12	291.13	292.99
Status	0.05	0.05	0.05	0.04	0.23	0.22	0.21	0.20
Credit line ($\times 100$)	93.05	91.24	94.34	93.56	41.13	39.10	36.78	38.52
Order amount ($\times 100$)	11.53	10.64	9.58	10.77	15.97	18.48	12.85	15.86
Borrow amount ($\times 100$)	8.58	9.01	13.26	13.69	15.86	15.51	18.83	21.88
Repay amount ($\times 100$)	6.83	6.95	12.69	11.37	13.93	12.44	18.49	19.20
Number of orders	3.82	3.87	4.09	4.55	3.51	3.37	2.85	4.46
Number of borrows	1.96	1.99	2.94	3.00	2.38	2.22	3.35	4.33
Number of repayment	0.97	0.94	1.13	1.11	0.51	0.58	0.53	0.68
Handling fee	15.87	16.47	14.09	16.29	63.26	67.60	51.66	59.90
Loan term	2.61	2.58	2.22	2.20	3.76	3.61	2.82	2.98
Monthly consumption	4.19	3.63	4.78	3.98	2.68	2.65	2.9	2.73

19.3% and $|0.49/0.61 - 1| = 19.7\%$ on in-sample and out-of-sample data, respectively. 2. For the Twins dataset, DRRB-OB has the second best result for ATE estimation, but still achieves the best result for ITE estimation, thereby showing the overall effectiveness. Note that DRRB-OB does not show significant priority on Twins experiments. This is probably because 1) the main objective for Twins data is classification (cross-entropy loss) instead of regression (mean squared loss); 2) unlike IHDP data where Y^0 and Y^1 are generated entirely by (\mathbf{X}, T) , Twins data exhibit a weaker connection between Y and (\mathbf{X}, T) since they are from randomized controlled trials, leading to underutilized strengths of our model as covariate balancing is less prone to affect outcome modeling in this case. 3. Compared to CFR-Wass, Dragonnet, and Dragon-Wass, DRRB shows consistent superiority, indicating that the proposed objective of DRRB-OB is constructive to treatment effect estimations.

4.2 Experiments on E-commerce Datasets: Credit

We introduce a new dataset for E-commerce credit policy evaluation, the **Credit** dataset. Traditional benchmark datasets such as IHDP generate ground truth of treatment effects through a parametric data generating form. Recently, Athey et al. (2021) have proposed using Wasserstein Generative Adversarial Networks (WGAN) (Arjovsky et al., 2017) to generate data with access to ground truth of treatment effects. The data generated by this nonparametric method approximately mimic the real-world data, making it more realistic than traditional parametric generating process.

Data description. The Credit dataset consists of 12-dimensional covariates in June, 2019: i) borrow amount, repayment amount, order amount, number of borrows, number of repayment, number of orders, handling fee, loan term, and credit line averaged in the past three months; ii) the default status of the last month; iii) age and days of usage. The treatment is the credit increase policy, and

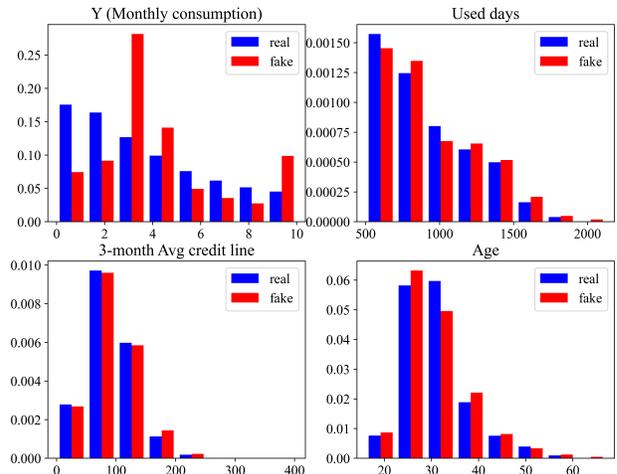


Figure 4: The distribution histogram for comparing the generated data and the real data with selected variables. In each graph, the x-axis represents the variable’s value, and the y-axis represents the probability density.

the outcome is the consumption amount within one month. The observational data exhibit selection bias because the assignment of the credit increase policy was determined by the E-commerce platform according to the consumers’ digital footprints (covariates). The original dataset contains 86880 individuals whose monthly consumption is less than 1000, of which 11056 are the treated units and 75824 are the controlled ones. Following the data generating algorithm in Athey et al. (2021), we apply WGAN to generate an artificial dataset containing 2000 samples whose factual and counterfactual outcomes are available. Then we compare the distribution of the generated data (fake data) with that of the original data (real data) in Table 2 and Figure 4. We defer more comparisons to Appendix.

Implementation. Following the strategy in creating IHDP datasets, we remove part of the treated units with the removal ratio varying in $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ for the generated Credit dataset. Then we randomly split the after-removal dataset by the ratio 56%/24%/20% as training/validation/test sets. We repeat this process 100 times to generate 100 different Credit datasets for each removal ratio. We remain the implementation details the same as used in the IHDP experiments, and defer more details of hyperparameters to Appendix. Furthermore, we make comparisons in Table 3 for DRRB, DRRB-OB, and other models that share very similar neural network architectures.

Results. As shown in Table 3, it is evident that DRRB and DRRB-OB can overwhelmingly achieve lower errors than other methods. Furthermore, we find DRRB-OB is better than DRRB in most cases, indicating the outcome balancing loss is effective. Moreover, it is worth noting that none of these baseline models can outperform TAR-

Table 3: Performance comparisons for with mean \pm standard error on out-of-sample **Credit** datasets. $\sqrt{\epsilon_{PEHE}}$: Lower is better. ϵ_{ATE} : Lower is better. Bold indicates the best results across different models.

Model / Removal ratio	Out-of-sample $\sqrt{\epsilon_{PEHE}}$					Out-of-sample ϵ_{ATE}				
	0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
TARNet	3.761 \pm .01	3.786 \pm .01	3.770 \pm .01	3.752 \pm .01	3.757 \pm .01	.280 \pm .03	.265 \pm .03	.265 \pm .03	.274 \pm .03	.277 \pm .03
CFR-Wass	3.766 \pm .01	3.801 \pm .01	3.773 \pm .01	3.768 \pm .01	3.761 \pm .01	.291 \pm .03	.270 \pm .03	.260 \pm .03	.317 \pm .04	.268 \pm .03
Dragonnet	3.769 \pm .01	3.791 \pm .01	3.771 \pm .01	3.750 \pm .01	3.755 \pm .01	.286 \pm .03	.282 \pm .03	.252 \pm .03	.280 \pm .03	.272 \pm .03
Dragon-Wass	3.759 \pm .01	3.784 \pm .01	3.771 \pm .01	3.752 \pm .01	3.753 \pm .01	.278 \pm .03	.271 \pm .03	.274 \pm .03	.276 \pm .03	.281 \pm .03
DRRB	3.758 \pm .01	3.783 \pm .01	3.771 \pm .01	3.747 \pm .01	3.748 \pm .01	.266 \pm .03	.253 \pm .03	.253 \pm .03	.265 \pm .03	.266 \pm .03
DRRB-OB	3.756 \pm .01	3.783 \pm .01	3.768 \pm .01	3.748 \pm .01	3.748 \pm .01	.260 \pm .03	.249 \pm .03	.230 \pm .02	.203 \pm .02	.262 \pm .03

Net (the objective of TARNet is equation (16)) across all the datasets. This reveals that it may be ineffectual if only considering domain invariance or domain discrimination. By contrast, DRRB and DRRB-OB exhibit stronger robustness and minor estimation errors than the baseline models, and such superiority is more significant in terms of ATE estimation. The experiments conducted on the Credit dataset suggest the DRRB-OB framework and the outcome balancing loss are practical and applicable to treatment effect estimations on real-world E-commerce credit data.

5 DISCUSSION

In this paper, we leverage balanced representation learning methods and design a new model, DRRB-OB, to do policy evaluation. In summary, the proposed method not only handles selection bias by learning domain-invariant representations, but also protects the learned representations from being over-balanced through a doubly robust loss. Moreover, we encourage representation balancing models to utilize the unconfoundedness assumption by incorporating the outcome balancing loss. Comprehensive experimental results on causal benchmark datasets and the Credit dataset confirm the effectiveness of our methods.

Limitation and future work.¹ Our study is limited to the PO framework, where the treatment is binary. Also, our method is not limited to the credit policy evaluation, and it can be extended to other problems where treatment effects need to be estimated. We believe another important question is how to design a proper adaptive strategy to determine the threshold in equation (19). Moreover, it would be interesting if future works go beyond the binary policy and explore further advances in balanced representation learning methods in terms of decision-making problems in the E-commerce field, e.g., the treatment assignment problem (Kitagawa and Tetenov, 2018; Athey and Wager, 2021). It is also worth noting that the consumers' credit loan is provided by the platform, so the platform is also exposed to credit risk. This prevents lenders from unethically allocating credit lines. Finally, it would be interesting if future commercial studies consider both consumption and credit

risk simultaneously. We hope our work can inspire more research on commercial credit policy under the representation learning framework, and we expect our work will boost more studies that intersect artificial intelligence, causal inference, and fintech to facilitate joint development in these areas.

Acknowledgements

Qi Wu acknowledges the support from The CityU-JD Digits Joint Laboratory in Financial Technology and Engineering; The Hong Kong Research Grants Council [General Research Fund 14206117, 11219420, and 11200219]; The CityU SRG-Fd fund 7005300, and The HK Institute of Data Science. The work described in this paper was partially supported by the InnoHK initiative, The Government of the HKSAR, and the Laboratory for AI-Powered Financial Technologies.

References

- Douglas Almond, Kenneth Y Chay, and David S Lee. The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083, 2005.
- Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017.
- Serge Assaad, Shuxi Zeng, Chenyang Tao, Shounak Datta, Nikhil Mehta, Ricardo Henao, Fan Li, and Lawrence Carin. Counterfactual representation learning with balancing weights. In *International Conference on Artificial Intelligence and Statistics*, pages 1972–1980. PMLR, 2021.
- Susan Athey and Stefan Wager. Policy learning with observational data. *Econometrica*, 89(1):133–161, 2021.
- Susan Athey, Guido W Imbens, Jonas Metzger, and Evan Munro. Using wasserstein generative adversarial networks for the design of monte carlo simulations. *Journal of Econometrics*, 2021.
- Shai Ben-David, John Blitzer, Koby Crammer, and Fernando Pereira. Analysis of representations for domain adaptation. *Advances in neural information processing systems*, 19, 2006.

¹We are thankful for all the reviewers and meta-reviewers who provide thoughtful and constructive comments and suggestions.

- Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine learning*, 79:151–175, 2010.
- Leo Breiman. Random forests. *Machine learning*, 45(1): 5–32, 2001.
- Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 2018.
- Hugh A Chipman, Edward I George, and Robert E McCulloch. Bart: Bayesian additive regression trees. *The Annals of Applied Statistics*, 4(1):266–298, 2010.
- Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*, 2015.
- Vincent Dorie. Nonparametric methods for causal inference. <https://github.com/vdorie/npci>, 2021.
- Xin Du, Lei Sun, Wouter Duivesteyn, Alexander Nikolaev, and Mykola Pechenizkiy. Adversarial balancing-based representation learning for causal effect inference with observational data. *Data Mining and Knowledge Discovery*, 35(4):1713–1738, 2021.
- Max H Farrell. Robust inference on average treatment effects with possibly more covariates than observations. *Journal of Econometrics*, 189(1):1–23, 2015.
- Ruocheng Guo, Jundong Li, Yichuan Li, K Selçuk Candan, Adrienne Raglin, and Huan Liu. Ignite: A minimax game toward learning individual treatment effects from networked observational data. In *IJCAI*, pages 4534–4540, 2020.
- James J Heckman and Edward Vytlacil. Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, 73(3):669–738, 2005.
- Jennifer L Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.
- Yiyan Huang, Cheuk Hang Leung, Xing Yan, Qi Wu, Nanbo Peng, Dongdong Wang, and Zhixiang Huang. The causal learning of retail delinquency. In *Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*, 2020.
- Yiyan Huang, Cheuk Hang Leung, Shumin Ma, Qi Wu, Dongdong Wang, and Zhixiang Huang. Moderately-balanced representation learning for treatment effects with orthogonality information. In *PRICAI 2022: Trends in Artificial Intelligence: 19th Pacific Rim International Conference on Artificial Intelligence, PRICAI 2022, Shanghai, China, November 10–13, 2022, Proceedings, Part II*, pages 3–16. Springer, 2022a.
- Yiyan Huang, Cheuk Hang Leung, Qi Wu, Xing Yan, Shumin Ma, Zhiri Yuan, Dongdong Wang, and Zhixiang Huang. Robust causal learning for the estimation of average treatment effects. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–9. IEEE, 2022b.
- Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *International conference on machine learning*, pages 3020–3029. PMLR, 2016.
- Yonghan Jung, Jin Tian, and Elias Bareinboim. Estimating identifiable causal effects through double machine learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 12113–12122, 2021.
- Edward H Kennedy. Optimal doubly robust estimation of heterogeneous causal effects. *arXiv preprint arXiv:2004.14497*, 2020.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Toru Kitagawa and Aleksey Tetenov. Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616, 2018.
- Michael C Knaus. A double machine learning approach to estimate the effects of musical practice on student’s skills. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 184(1):282–300, 2021.
- Fan Li, Kari Lock Morgan, and Alan M Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.
- Molei Liu, Yi Zhang, and Doudou Zhou. Double/debiased machine learning for logistic partially linear model. *The Econometrics Journal*, 24(3):559–588, 2021.
- Christos Louizos, Uri Shalit, Joris Mooij, David Sontag, Richard Zemel, and Max Welling. Causal effect inference with deep latent-variable models. *arXiv preprint arXiv:1705.08821*, 2017.
- Jared K Lunceford and Marie Davidian. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in medicine*, 23(19):2937–2960, 2004.
- Miruna Oprescu, Vasilis Syrgkanis, and Zhiwei Steven Wu. Orthogonal random forest for causal inference. In *International Conference on Machine Learning*, pages 4932–4941. PMLR, 2019.
- Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.
- Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the*

American Statistical Association, 100(469):322–331, 2005.

Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International Conference on Machine Learning*, pages 3076–3085. PMLR, 2017.

Claudia Shi, David Blei, and Victor Veitch. Adapting neural networks for the estimation of treatment effects. *Advances in neural information processing systems*, 32, 2019.

Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Jui-Chung Yang, Hui-Ching Chuang, and Chung-Ming Kuan. Double machine learning with gradient boosting and its application to the big n audit quality effect. *Journal of Econometrics*, 216(1):268–283, 2020.

Liuyi Yao, Sheng Li, Yaliang Li, Mengdi Huai, Jing Gao, and Aidong Zhang. Representation learning for treatment effect estimation from observational data. *Advances in Neural Information Processing Systems*, 31, 2018.

Jinsung Yoon, James Jordon, and Mihaela Van Der Schaar. Ganite: Estimation of individualized treatment effects using generative adversarial nets. In *International Conference on Learning Representations*, 2018.

Yao Zhang, Alexis Bellot, and Mihaela Schaar. Learning overlapping representations for the estimation of individualized treatment effects. In *International Conference on Artificial Intelligence and Statistics*, pages 1005–1014. PMLR, 2020.

A Proofs of Technical Results

To prove theorem 1, we first introduce some necessary terms.

Definition 1. Let Φ be the representation function such that $\Phi : \mathcal{X} \rightarrow \mathcal{R}$, $h : \mathcal{R} \times \{0, 1\} \rightarrow \mathcal{Y}$ be the outcome function such that $h(\Phi(\mathbf{x}), t)$ estimates y^t , and $\ell_{h, \Phi}(\mathbf{x}, t) = \int_{\mathcal{Y}} (y^t - h(\Phi(\mathbf{x}), t))^2 p(y^t | \mathbf{x}) dy^t$. The expected factual loss $\epsilon_F(h, \Phi)$ and counterfactual losses $\epsilon_{CF}(h, \Phi)$ over h and Φ are, respectively:

$$\begin{aligned}\epsilon_F(h, \Phi) &= \int_{\mathcal{X} \times \{0, 1\}} \ell_{h, \Phi}(\mathbf{x}, t) p(\mathbf{x}, t) d\mathbf{x} dt, \\ \epsilon_{CF}(h, \Phi) &= \int_{\mathcal{X} \times \{0, 1\}} \ell_{h, \Phi}(\mathbf{x}, t) p(\mathbf{x}, 1 - t) d\mathbf{x} dt.\end{aligned}$$

Lemma 2. Let $\Phi : \mathcal{X} \rightarrow \mathcal{R}$ be an invertible representation with Ψ being its inverse. Let $p_{\Phi}^{T=1}$ and $p_{\Phi}^{T=0}$ be the densities of the treated and controlled covariates in the representation space. Let $h : \mathcal{R} \times \{0, 1\} \rightarrow \mathcal{Y}$, $u := Pr(T = 1)$ and \mathcal{G} be the family of 1-Lipschitz functions. Assume there exists a constant $B_{\Phi} \geq 0$, such that for $t = 0, 1$, the function $g_{\Phi, h}(\mathbf{r}, t) := \frac{1}{B_{\Phi}} \cdot \ell_{h, \Phi}(\Psi(\mathbf{r}), t) \in \mathcal{G}$. Then we have:

$$\epsilon_{CF}(h, \Phi) \leq (1 - u) \cdot \epsilon_F^{T=1}(h, \Phi) + u \cdot \epsilon_F^{T=0}(h, \Phi) + B_{\Phi} \cdot \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}).$$

Proof.

$$\begin{aligned}& \epsilon_{CF}(h, \Phi) - [(1 - u) \cdot \epsilon_F^{T=1}(h, \Phi) + u \cdot \epsilon_F^{T=0}(h, \Phi)] \\ &= [(1 - u) \cdot \epsilon_{CF}^{T=1}(h, \Phi) + u \cdot \epsilon_{CF}^{T=0}(h, \Phi)] - [(1 - u) \cdot \epsilon_F^{T=1}(h, \Phi) + u \cdot \epsilon_F^{T=0}(h, \Phi)] \\ &= (1 - u) \cdot [\epsilon_{CF}^{T=1}(h, \Phi) - \epsilon_F^{T=1}(h, \Phi)] + u \cdot [\epsilon_{CF}^{T=0}(h, \Phi) - \epsilon_F^{T=0}(h, \Phi)] \\ &= (1 - u) \int_{\mathcal{X}} \ell_{h, \Phi}(\mathbf{x}, 1) (p^{T=0}(\mathbf{x}) - p^{T=1}(\mathbf{x})) d\mathbf{x} + u \int_{\mathcal{X}} \ell_{h, \Phi}(\mathbf{x}, 0) (p^{T=1}(\mathbf{x}) - p^{T=0}(\mathbf{x})) d\mathbf{x} \\ &= (1 - u) \int_{\mathcal{R}} \ell_{h, \Phi}(\Psi(\mathbf{r}), 1) (p_{\Phi}^{T=0}(\mathbf{r}) - p_{\Phi}^{T=1}(\mathbf{r})) d\mathbf{r} + u \int_{\mathcal{R}} \ell_{h, \Phi}(\Psi(\mathbf{r}), 0) (p_{\Phi}^{T=1}(\mathbf{r}) - p_{\Phi}^{T=0}(\mathbf{r})) d\mathbf{r} \\ &\leq B_{\Phi} \cdot (1 - u) \int_{\mathcal{R}} \frac{1}{B_{\Phi}} \ell_{h, \Phi}(\Psi(\mathbf{r}), 1) (p_{\Phi}^{T=0}(\mathbf{r}) - p_{\Phi}^{T=1}(\mathbf{r})) d\mathbf{r} \\ &\quad + B_{\Phi} \cdot u \int_{\mathcal{R}} \frac{1}{B_{\Phi}} \ell_{h, \Phi}(\Psi(\mathbf{r}), 0) (p_{\Phi}^{T=1}(\mathbf{r}) - p_{\Phi}^{T=0}(\mathbf{r})) d\mathbf{r} \\ &\leq B_{\Phi} \cdot (1 - u) \sup_{g \in \mathcal{G}} \left| \int_{\mathcal{R}} g(\mathbf{r}) (p_{\Phi}^{T=0}(\mathbf{r}) - p_{\Phi}^{T=1}(\mathbf{r})) d\mathbf{r} \right| \\ &\quad + B_{\Phi} \cdot u \cdot \sup_{g \in \mathcal{G}} \left| \int_{\mathcal{R}} g(\mathbf{r}) (p_{\Phi}^{T=1}(\mathbf{r}) - p_{\Phi}^{T=0}(\mathbf{r})) d\mathbf{r} \right| \\ &= B_{\Phi} \cdot \sup_{g \in \mathcal{G}} \left| \int_{\mathcal{R}} g(\mathbf{r}) (p_{\Phi}^{T=1}(\mathbf{r}) - p_{\Phi}^{T=0}(\mathbf{r})) d\mathbf{r} \right| \\ &= B_{\Phi} \cdot \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}).\end{aligned} \tag{21}$$

□

Lemma 3. Let loss function L be the squared loss such that $L(y_1, y_2) = (y_1 - y_2)^2$. Defining $\sigma_y^2 = \min\{\sigma_{y^t}^2(p(\mathbf{x}, t)), \sigma_{y^t}^2(p(\mathbf{x}, 1 - t))\} \forall t \in \{0, 1\}$, where $\sigma_{y^t}^2(p(\mathbf{x}, t)) = \int_{\mathcal{X} \times \{0, 1\} \times \mathcal{Y}} (y^t - \tau^t(\mathbf{x}))^2 p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt$, we have

$$\epsilon_{PEHE}(h, \Phi) \leq 2(\epsilon_{CF}(h, \Phi) + \epsilon_F(h, \Phi) - 2\sigma_y^2).$$

Proof. Define the function $f : \mathcal{X} \times \{0, 1\} \rightarrow \mathcal{Y}$ such that $f(\mathbf{x}, t) = h(\Phi(\mathbf{x}), t)$. We denote $\epsilon_{PEHE}(f) = \epsilon_{PEHE}(h, \Phi)$, $\epsilon_F(f) = \epsilon_F(h, \Phi)$, $\epsilon_{CF}(f) = \epsilon_{CF}(h, \Phi)$ for $f(\mathbf{x}, t) = h(\Phi(\mathbf{x}), t)$. Assuming that $\tau^1(\mathbf{x}) := \mathbb{E}[Y^1 | \mathbf{X} = \mathbf{x}]$ and $\tau^0(\mathbf{x}) :=$

$\mathbb{E}[Y^0 | \mathbf{X} = \mathbf{x}]$, then we have

$$\begin{aligned}
 \epsilon_{PEHE}(f) &= \int_{\mathcal{X}} ((f(\mathbf{x}, 1) - f(\mathbf{x}, 0)) - (\tau^1(\mathbf{x}) - \tau^0(\mathbf{x})))^2 p(\mathbf{x}) d\mathbf{x} \\
 &\leq 2 \int_{\mathcal{X}} ((f(\mathbf{x}, 1) - \tau^1(\mathbf{x}))^2 + (f(\mathbf{x}, 0) - \tau^0(\mathbf{x}))^2) p(\mathbf{x}) d\mathbf{x} \\
 &= 2 \int_{\mathcal{X}} (f(\mathbf{x}, 1) - \tau^1(\mathbf{x}))^2 p(\mathbf{x}, T=1) d\mathbf{x} + 2 \int_{\mathcal{X}} (f(\mathbf{x}, 0) - \tau^0(\mathbf{x}))^2 p(\mathbf{x}, T=0) d\mathbf{x} \\
 &\quad + 2 \int_{\mathcal{X}} (f(\mathbf{x}, 1) - \tau^1(\mathbf{x}))^2 p(\mathbf{x}, T=0) d\mathbf{x} + 2 \int_{\mathcal{X}} (f(\mathbf{x}, 0) - \tau^0(\mathbf{x}))^2 p(\mathbf{x}, T=1) d\mathbf{x} \\
 &= 2 \int_{\mathcal{X} \times \{0,1\}} (f(\mathbf{x}, t) - \tau^t(\mathbf{x}))^2 p(\mathbf{x}, t) d\mathbf{x} dt + 2 \int_{\mathcal{X} \times \{0,1\}} (f(\mathbf{x}, t) - \tau^t(\mathbf{x}))^2 p(\mathbf{x}, 1-t) d\mathbf{x} dt.
 \end{aligned} \tag{22}$$

On the other hand,

$$\begin{aligned}
 \epsilon_F(f) &= \int_{\mathcal{X} \times \{0,1\} \times \mathcal{Y}} (f(\mathbf{x}, t) - y^t)^2 p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt \\
 &= \int_{\mathcal{X} \times \{0,1\} \times \mathcal{Y}} (f(\mathbf{x}, t) - \tau^t(\mathbf{x}))^2 p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt \\
 &\quad + \int_{\mathcal{X} \times \{0,1\} \times \mathcal{Y}} (\tau^t(\mathbf{x}) - y^t)^2 p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt \\
 &\quad + 2 \int_{\mathcal{X} \times \{0,1\} \times \mathcal{Y}} (f(\mathbf{x}, t) - \tau^t(\mathbf{x})) (\tau^t(\mathbf{x}) - y^t) p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt \\
 &= \int_{\mathcal{X} \times \{0,1\}} (f(\mathbf{x}, t) - \tau^t(\mathbf{x}))^2 p(\mathbf{x}, t) d\mathbf{x} dt + \sigma_{y^t}^2(p(\mathbf{x}, t))
 \end{aligned} \tag{23}$$

Similarly, for ϵ_{CF} , we have

$$\epsilon_{CF}(f) = \int_{\mathcal{X} \times \{0,1\}} (f(\mathbf{x}, t) - \tau^t(\mathbf{x}))^2 p(\mathbf{x}, 1-t) d\mathbf{x} dt + \sigma_{y^t}^2(p(\mathbf{x}, 1-t)). \tag{24}$$

Combining results (22), (23), and (24), we have

$$\begin{aligned}
 \epsilon_{PEHE}(h, \Phi) &\leq 2(\epsilon_F(f) - \sigma_{y^t}^2(p(\mathbf{x}, t))) + 2(\epsilon_{CF}(f) - \sigma_{y^t}^2(p(\mathbf{x}, 1-t))) \\
 &\leq 2(\epsilon_{CF}(h, \Phi) + \epsilon_F(h, \Phi) - 2\sigma_y^2).
 \end{aligned} \tag{25}$$

□

Theorem 4. Let Ψ be the inverse of Φ , $p_{\Phi}^{T=1}$ and $p_{\Phi}^{T=0}$ be the densities of the treated and controlled covariates in the representation space, $h : \mathcal{R} \times \{0, 1\} \rightarrow \mathcal{Y}$ such that $h(\Phi(\mathbf{x}), t)$ estimates y^t , and G be the family of 1-Lipschitz functions. Assume that there exists a constant $B_{\Phi} \geq 0$ such that $g_{\Phi, h}(\mathbf{r}, t) := \frac{1}{B_{\Phi}} \cdot \ell_{h, \Phi}(\Psi(\mathbf{r}), t) \in G$ for any $t \in \{0, 1\}$, where $\ell_{h, \Phi}(\mathbf{x}, t) = \int_{\mathcal{Y}} (y^t, h(\Phi(\mathbf{x}), t))^2 p(y^t | \mathbf{x}) dy^t$. Defining $\sigma_y^2 = \min\{\sigma_{y^t}^2(p(\mathbf{x}, t)), \sigma_{y^t}^2(p(\mathbf{x}, 1-t))\} \forall t \in \{0, 1\}$, where $\sigma_{y^t}^2(p(\mathbf{x}, t)) = \int_{\mathcal{X} \times \{0,1\} \times \mathcal{Y}} (y^t - \tau^t(\mathbf{x}))^2 p(y^t | \mathbf{x}) p(\mathbf{x}, t) dy^t d\mathbf{x} dt$, we then have

$$\begin{aligned}
 \epsilon_{PEHE}(h, \Phi) &\leq 2(\epsilon_F^{T=1}(h, \Phi) + \epsilon_F^{T=0}(h, \Phi) \\
 &\quad + B_{\Phi} \cdot \text{IPM}_G(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}) - 2\sigma_y^2).
 \end{aligned} \tag{26}$$

Proof. By Lemma 2, Lemma 3 and the fact that $\epsilon_F(h, \Phi) = u \cdot \epsilon_F^{T=1}(h, \Phi) + (1-u) \cdot \epsilon_F^{T=0}(h, \Phi)$, we have

$$\begin{aligned}
 \epsilon_{PEHE}(h, \Phi) &\leq 2(\epsilon_{CF}(h, \Phi) + \epsilon_F(h, \Phi) - 2\sigma_y^2) \\
 &\leq 2((1-u) \cdot \epsilon_F^{T=1}(h, \Phi) + u \cdot \epsilon_F^{T=0}(h, \Phi) + B_{\Phi} \cdot \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}) + \epsilon_F(h, \Phi) - 2\sigma_y^2) \\
 &= 2((1-u) \cdot \epsilon_F^{T=1}(h, \Phi) + u \cdot \epsilon_F^{T=0}(h, \Phi) + B_{\Phi} \cdot \text{Wass}(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}) \\
 &\quad + u \cdot \epsilon_F^{T=1}(h, \Phi) + (1-u) \cdot \epsilon_F^{T=0}(h, \Phi) - 2\sigma_y^2) \\
 &= 2(\epsilon_F^{T=1}(h, \Phi) + \epsilon_F^{T=0}(h, \Phi) + B_{\Phi} \cdot \text{IPM}_G(p_{\Phi}^{T=1}, p_{\Phi}^{T=0}) - 2\sigma_y^2).
 \end{aligned} \tag{27}$$

□

B Additional Experimental Results

Results. We report performances produced by our model DRRB-OB and other baseline models in Table 1. The symbol “*” means that their paper does not report this result. In addition, we plot the loss outputs for the first 100 iterations of RB, DRRB and DRRB-OB on the 1st IHDP dataset in Figure 5. Figure 5 shows that RB suffers the over-balancing issue (with W_{ass} being very close to 0), and hence ϵ_F does not get well minimized. As a result, $\sqrt{\epsilon_{PEHE}}$ converges to the value around 1.0 after 25 iterations. In contrast, DRRB and DRRB-OB can achieve smaller ϵ_F with the iteration increasing, thereby leading to smaller $\sqrt{\epsilon_{PEHE}}$ with the value around 0.4. Additionally, it should be noticed that the optimization of DRRB-OB is more steady than that of DRRB.

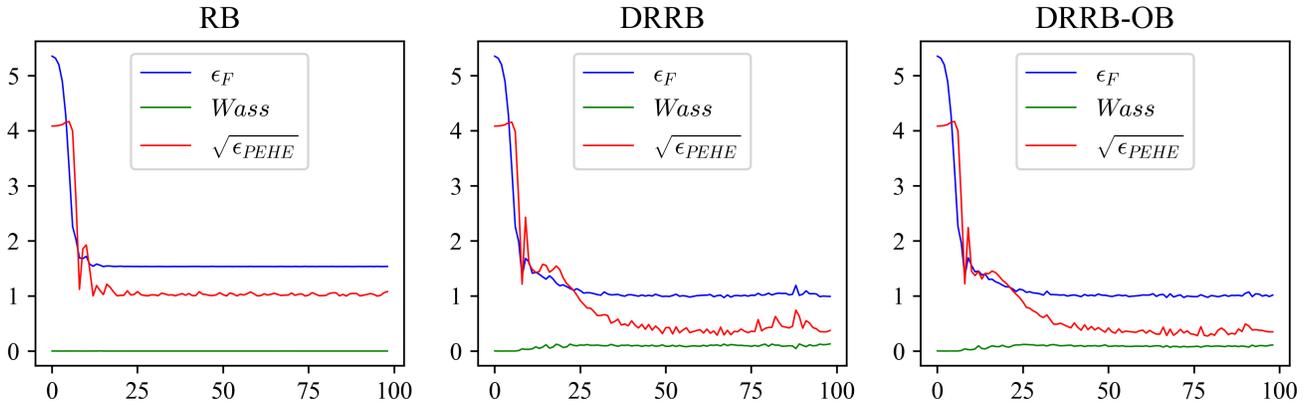


Figure 5: Loss outputs for the first 100 iterations of RB, DRRB, and DRRB-OB.

Comparisons between fake and real data We plot distribution histograms for comparing the generated (fake) data and the original (real) data in Figure 6. This illustrates that using the strategy in Athey et al. (2021) to generate monte carlo simulations is reasonable since the generated data are similarly distributed to the original data.

Implementation details We set 5/3/3/3 fully connected layers with 100/100/100/100 units for the network $\Phi/\pi/f_0/f_1$ on benchmark experiments, and set 4/3/3/3 fully connected layers with 200/100/200/200 units for the network $\Phi/\pi/f_0/f_1$ on Credit experiments. All experiments adopt the activation function as ELU (Clevert et al., 2015). For IHDP/Twins/Credit experiments, the (batch size, epoch) are set to be (100, 600)/(1000, 300)/(200, 200), and (λ_1, λ_2) are set to be (0.1, 1)/(0.01, 1)/(0.1, 30). Note that λ_2 is chosen according to the doubly robust loss on the validation set within the first 10 iterations, and it can also be determined adaptively, i.e., updated in each iteration. The optimizer is chosen as Adam (Kingma and Ba, 2014), and the learning rates for the three optimization steps are set to be $1e^{-3}/1e^{-3}/1e^{-3}$, $1e^{-3}/1e^{-4}/1e^{-3}$, $1e^{-3}/1e^{-4}/1e^{-3}$ for IHDP, Twins, Credit experiments, respectively. All methods are selected by the searching rule and early stopping rule stated in Shalit et al. (2017), and the hyperparameters of baseline models are consistent with their papers or public codes.

Table 4: Performance comparisons with mean \pm standard error on benchmark datasets **IHDP** and **Twins**. $\sqrt{\epsilon_{PEHE}}$: Lower is better. ϵ_{ATE} : Lower is better. AUC: Higher is better.

IHDP				
Method	Within-sample		Out-of-sample	
	$\sqrt{\epsilon_{PEHE}}$	ϵ_{ATE}	$\sqrt{\epsilon_{PEHE}}$	ϵ_{ATE}
OLS/LR ₁	5.8 \pm .3	.73 \pm .04	5.8 \pm .3	.94 \pm .06
OLS/LR ₂	2.4 \pm .1	.14 \pm .01	2.5 \pm .1	.31 \pm .02
k-NN	2.1 \pm .1	.14 \pm .01	4.1 \pm .2	.79 \pm .05
BART Chipman et al. (2010)	2.1 \pm .1	.23 \pm .01	2.3 \pm .1	.34 \pm .02
RF Breiman (2001)	4.2 \pm .2	.73 \pm .05	6.6 \pm .3	.96 \pm .06
CF Wager and Athey (2018)	3.8 \pm .2	.18 \pm .01	3.8 \pm .2	.40 \pm .03
CEVAE Louizos et al. (2017)	2.7 \pm .1	.34 \pm .01	2.6 \pm .1	.46 \pm .02
SITE Yao et al. (2018)	.69 \pm .0	.22 \pm .01	.75 \pm .0	.24 \pm .01
GANITE Yoon et al. (2018)	1.9 \pm .4	.43 \pm .05	2.4 \pm .4	.49 \pm .05
BLR Johansson et al. (2016)	5.8 \pm .3	.72 \pm .04	5.8 \pm .3	.93 \pm .05
BNN Johansson et al. (2016)	2.2 \pm .1	.37 \pm .03	2.1 \pm .1	.42 \pm .03
TARNet Shalit et al. (2017)	.88 \pm .0	.26 \pm .01	.95 \pm .0	.28 \pm .01
CFR-WASS Shalit et al. (2017)	.71 \pm .0	.25 \pm .01	.76 \pm .0	.27 \pm .01
Dragonnet Shi et al. (2019)	.89 \pm .0	.15 \pm .01	.93 \pm .0	.18 \pm .01
DKLITE Zhang et al. (2020)	.52 \pm .0	*	.65 \pm .0	*
BWCFR-MW Assaad et al. (2021)	*	*	.66 \pm .0	.18 \pm .01
BWCFR-OW Assaad et al. (2021)	*	*	.65 \pm .1	.18 \pm .02
BWCFR-TruncIPW Assaad et al. (2021)	*	*	.63 \pm .0	.19 \pm .01
ABCEI Du et al. (2021)	.96 \pm .0	.18 \pm .01	1.1 \pm .0	.20 \pm .01
Dragon-Wass	.82 \pm .0	.14 \pm .01	.86 \pm .0	.17 \pm .01
DRRB-OB (Ours)	.46 \pm .0	.12 \pm .01	.49 \pm .0	.13 \pm .01

Twins				
Method	Within-sample		Out-of-sample	
	AUC	ϵ_{ATE}	AUC	ϵ_{ATE}
OLS/LR ₁	.660 \pm .005	.004 \pm .003	.500 \pm .028	.007 \pm .006
OLS/LR ₂	.660 \pm .004	.004 \pm .003	.500 \pm .016	.007 \pm .006
k-NN	.609 \pm .010	.003 \pm .002	.492 \pm .012	.005 \pm .004
BART Chipman et al. (2010)	.506 \pm .014	.121 \pm .024	.500 \pm .011	.127 \pm .024
RF Breiman (2001)	*	.005 \pm .003	*	.008 \pm .005
CF Wager and Athey (2018)	*	.029 \pm .004	*	.034 \pm .008
CEVAE Louizos et al. (2017)	.845 \pm .003	.022 \pm .002	.841 \pm .004	.032 \pm .003
SITE Yao et al. (2018)	.862 \pm .002	.016 \pm .001	.853 \pm .006	.020 \pm .002
GANITE Yoon et al. (2018)	*	.006 \pm .002	*	.009 \pm .008
BLR Johansson et al. (2016)	.611 \pm .009	.006 \pm .004	.510 \pm .018	.033 \pm .009
BNN Johansson et al. (2016)	.690 \pm .008	.006 \pm .003	.676 \pm .008	.020 \pm .007
TARNet Shalit et al. (2017)	.849 \pm .002	.011 \pm .002	.840 \pm .006	.015 \pm .002
CFR-WASS Shalit et al. (2017)	.850 \pm .002	.011 \pm .002	.842 \pm .005	.028 \pm .003
Dragonnet Shi et al. (2019)	.880 \pm .000	.005 \pm .005	.874 \pm .001	.008 \pm .005
ABCEI Du et al. (2021)	.871 \pm .001	.003 \pm .001	.863 \pm .001	.005 \pm .001
Dragon-Wass	.877 \pm .000	.004 \pm .001	.874 \pm .001	.007 \pm .001
DRRB-OB (Ours)	.880 \pm .000	.004 \pm .000	.875 \pm .001	.008 \pm .001

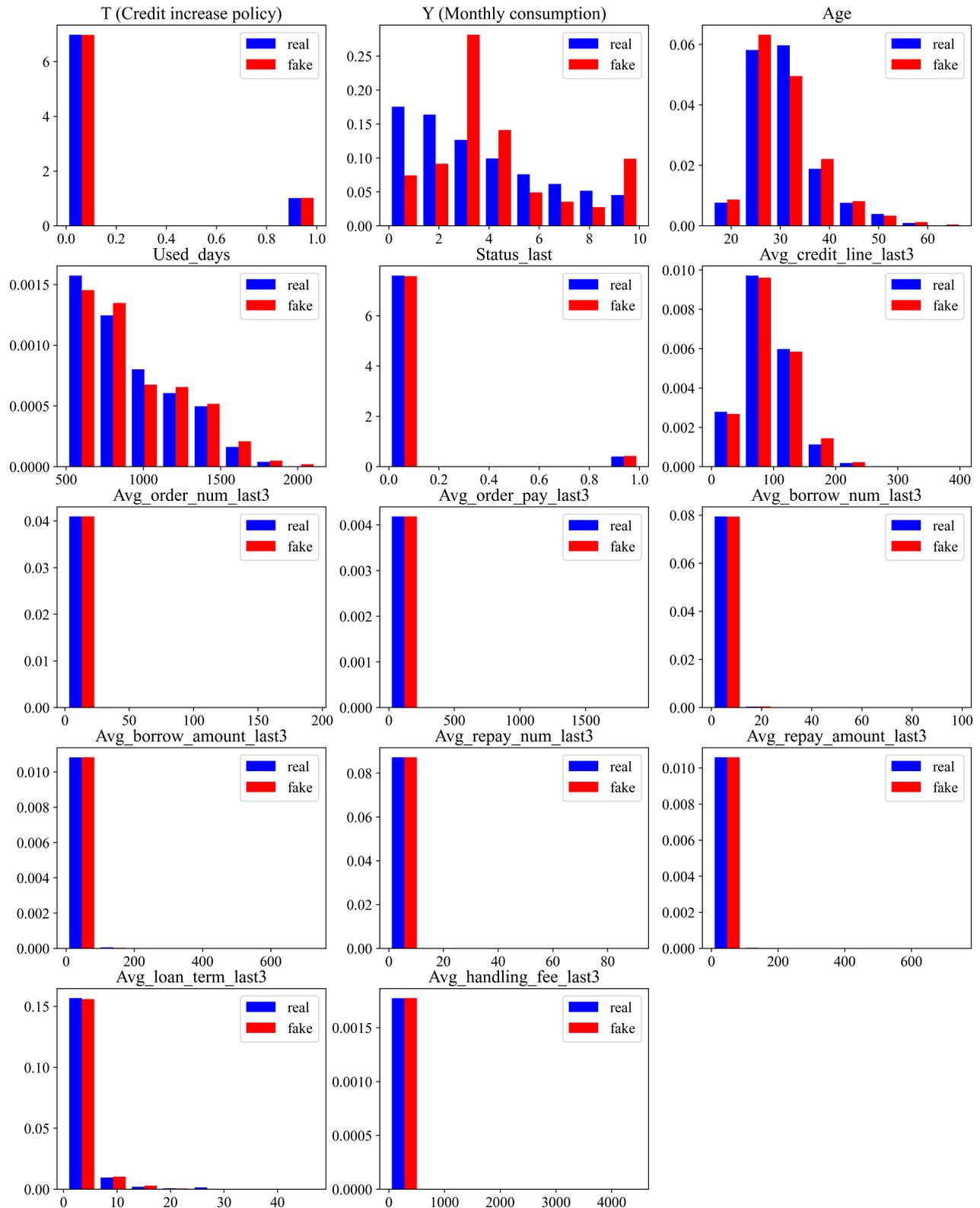


Figure 6: The distribution histogram for comparing the generated data and the real data with selected variables. In each graph, the x-axis represents the variable's value, and the y-axis represents the probability density. The histogram is plotted using `matplotlib.pyplot.hist(real, fake, density=True)`.