Learning Conditional Instrumental Variable Representation for Causal Effect Estimation

Debo Cheng$^{(\boxtimes)}$, Ziqi Xu, Jiuyong Li, Lin Liu, Thuc Duy Le, and Jixue Liu

University of South Australia, Adelaide, Australia
{Debo.Cheng,Ziqi.Xu}@mymail.unisa.edu.au,
{Jiuyong.Li,Lin.Liu,Thuc.Le,Jixue.Liu}@unisa.edu.au

Abstract. One of the fundamental challenges in causal inference is to estimate the causal effect of a treatment on its outcome of interest from observational data. However, causal effect estimation often suffers from the impacts of confounding bias caused by unmeasured confounders that affect both the treatment and the outcome. The instrumental variable (IV) approach is a powerful way to eliminate the confounding bias from latent confounders. However, the existing IV-based estimators require a nominated IV, and for a conditional IV (CIV) the corresponding conditioning set too, for causal effect estimation. This limits the application of IV-based estimators. In this paper, by leveraging the advantage of disentangled representation learning, we propose a novel method, named DVAE.CIV, for learning and disentangling the representations of CIV and the representations of its conditioning set for causal effect estimations from data with latent confounders. Extensive experimental results on both synthetic and real-world datasets demonstrate the superiority of the proposed DVAE.CIV method against the existing causal effect estimators.

Keywords: Causal Inference · Instrumental Variable · Latent Confounder

1 Introduction

It is a fundamental task to query or estimate the causal effect of a treatment (a.k.a. exposure, intervention or action) on an outcome of interest in causal inference. Causal effect estimation has wide applications across many fields, including but not limited to, economics [19], epidemiology [16,28], and computer science [30]. The gold standard method for causal effect estimation is randomised controlled trials (RCT), but they are often impractical or unethical due to cost restrictions or ethical constraints [19,30]. Instead of conducting an RCT, estimating causal effects from observational data offers an alternative to evaluate the effect of a treatment on the outcome of interest.

D. Cheng and Z. Xu—These authors contributed equally.
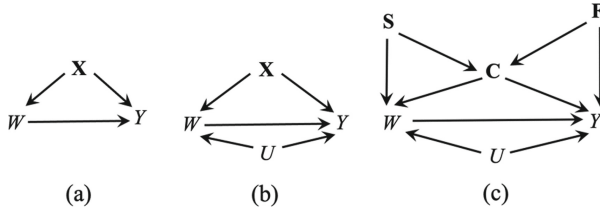
Fig. 1. Three causal DAGs are utilised to illustrate the problems of causal effect estimation from observational data. In all three DAGs, $\mathbf{X}$, U, W and Y are the set of pretreatment variables, latent confounder, treatment and outcome variables, respectively. (a) indicates the unconfoundedness assumption holding, and (b) shows the causal effect of W on Y is non-identification since there is a latent confounder U. (c) illustrates the problem studied in this work, in which the set $\mathbf{X}$ is represented by three sets $\{\mathbf{S}, \mathbf{C}, \mathbf{F}\}$.

Confounding bias is a major obstacle in estimating causal effects from observational data. It arises from confounders that affect both the treatment variable W and the outcome variable Y. When all confounders are measured (i.e., the unconfoundedness assumption [19,31] is satisfied), adjusting for the set of all measured confounders is sufficient to obtain an unbiased estimation of the causal effect from observational data [1,19]. For example, in the causal graph of Fig. 1(a), the unconfoundedness is satisfied when given $\mathbf{X}$. Nevertheless, the unconfoundedness assumption is untestable, and there exists a latent (a.k.a. unobserved, unmeasured) confounder affecting both W and Y in many real-world applications, e.g. the latent confounder U affects both W and Y in the causal graph in Fig. 1(b). In such a situation, the causal effect of W on Y is non-identification [30]. Most existing data-driven methods rely on the unconfoundedness assumption and thus it becomes challenging and questionable for them to obtain unbiased causal effects from data with latent confounders.

The instrumental variable (IV) approach is a practical and powerful technique for addressing the challenging problem of causal effect estimation in the presence of latent confounders. The IV approach requires a valid IV for eliminating the confounding bias caused by latent confounders [2,18]. Valid IVs are exogenous variables that are associated with W but not directly associated with Y [16,27]. A valid IV S needs to satisfy three conditions: (1) S is correlated to W; (2) S and Y do not share confounders (i.e. *unconfounded instrument*); and (3) the effect of S on Y is entirely through W (i.e. *exogenous*) [16,27]. However, the last two conditions are too strict and not testable in real-world applications. Therefore, in many existing IV-based methods, an IV is nominated based on prior or domain knowledge. However, in many real-world applications, the nominated IVs based on domain knowledge could violate one of the three conditions, resulting in a biased estimate and potentially leading to incorrect conclusions [6,16].

It is a challenging problem to discover a valid IV directly from data. Investigators usually collect as many covariates as possible, but few of them are valid IVs that satisfy the three conditions. Instead of discovering a valid IV, Kang et al. [20] proposed a data-driven method, referred to as sisVIVE, based on the assumption of some invalid and some valid IVs (i.e. more than half of candidate IVs are valid IVs) to provide a bound of causal effect estimations. Hartford et al. [14] proposed DeepIV, a deep learning based IV approach for counterfactual predictions, but it requires a nominated IV and the corresponding conditioning set. Kuang et al. [23] developed a method to model a summary IV as a latent variable based on the statistical dependencies of the set of candidate IVs. Yuan et al. [36] proposed a data-driven method to automatically generate a synthetic IV for counterfactual predictions, but the method does not consider the confounding bias between the IV and the outcome, and the condition of *unconfounded instrument* may be violated in many cases. Therefore, it is desirable to develop an algorithm for learning a valid IV that considers the unconfounded instrument for causal effect estimations, especially conditional average causal effect estimations, from data with latent confounders.

To provide a practical solution for conditional average causal effect estimations, in this work, we focus on conditional IV (CIV), which can be considered as an IV with relaxed conditions and a CIV requires a conditioning set to instrumentalise it to function as an IV (details see Definition 1). We propose to leverage disentangled representation learning technique to learn from data the representations of a CIV and its conditioning set.

Specifically, as shown in Fig. 1(c), we assume that the observed covariates are learned through three representations, $\mathbf{S}$, $\mathbf{C}$ and $\mathbf{F}$. Here, $\mathbf{S}$ affects both treatment W and $\mathbf{C}$, $\mathbf{C}$ represents the confounding factor affecting both W and the outcome Y, and $\mathbf{F}$ represents the risk factor affecting both $\mathbf{C}$ and Y. We then establish a theorem that $\mathbf{S}$ is a valid CIV that is instrumentalised by $\{\mathbf{C}, \mathbf{F}\}$, meaning that $\{\mathbf{C}, \mathbf{F}\}$ is the conditioning set of $\mathbf{S}$. Supported by this theorem, we design and develop a novel disentangled representation learning algorithm called DVAE.CIV model, which is based on the Variational AutoEncoder (VAE) model [22]. This model allows us to obtain the representations of the CIV $\mathbf{S}$ and its conditioning set $\{\mathbf{C}, \mathbf{F}\}$, enabling us to use $\mathbf{S}$ as a valid IV conditioning on $\{\mathbf{C}, \mathbf{F}\}$ for estimating the conditional average causal effects of W on Y from data when there are latent confounders. The main contributions of the paper are summarised as follows.

– We address a challenging problem in conditional average causal effect estimations from data with latent confounders by utilising the CIV approach and VAE models.
– We propose a novel disentanglement learning model based on the conditional VAE model to learn and disentangle the representations of covariates into the representations of a CIV $\mathbf{S}$ and its conditioning set $\{\mathbf{C}, \mathbf{F}\}$ for conditional average causal effect estimations from data with latent confounders.

– We conduct extensive experiments on synthetic and real-world datasets to show the performance of the DVAE.CIV model, w.r.t. causal effect estimations from data with latent confounders.

2   Preliminaries

In this paper, uppercase and lowercase letters are utilised to represent variables and their values, respectively. Bold-faced uppercase and lowercase letters indicate a set of variables and a value assignment of the set, respectively.

A DAG (direct acyclic graph) is a graph that contains directed edges (i.e. $\rightarrow$) without cycles. In a DAG $\mathcal{G}$, the directed edge $X_i \rightarrow X_j$ represents that X_i is a cause of X_j, and X_j is an effect of X_i. A DAG is a causal DAG when a direct edge $X_i \rightarrow X_j$ represents that X_i is a cause of X_j. In this work, we assume a causal DAG $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ to represent the underlying system, where $\mathbf{V} = \mathbf{X} \cup \mathbf{U} \cup \{W, Y\}$, and $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$ denotes directed edges. In $\mathbf{V}$, we assume that $\mathbf{X}$ is the set of pretreatment variables, $\mathbf{U}$ is the set of latent confounders, W is a binary treatment variable ($w = 1$ and $w = 0$ denote the treated sample and control sample, respectively), and $Y(w)$ is an outcome of interest. Following the potential outcome model [19,31], we have the potential outcomes $Y(w = 1)$ and $Y(w = 0)$ relative to the treatment W. Note that we can only measure one of the two potential outcomes for a given individual x_i. Conceptually, the individual causal effect (ICE) at x_i is defined as $ICE_i = Y_i(w = 1) - Y_i(w = 0)$. The average causal effect of W on Y is defined as $\mathrm{ACE}(W, Y) = \mathbb{E}[Y_i(w = 1) - Y_i(w = 0)]$, where $\mathbb{E}$ is the expectation function.

The conditional average causal effect (CACE) of W on Y is referred to as $\mathrm{CACE}(W, Y)$, and defined as the form $P(Y|do(w), \mathbf{X})$, where $do(\cdot)$ is do-operation and indicates an intervention on the treatment (i.e. set the value of W as per [30]). Conceptually, $P(Y|do(w), \mathbf{X})$ can be obtained as:

$$\mathrm{CACE}(W, Y) = \mathbb{E}[Y_i(w = 1) - Y_i(w = 0) \mid \mathbf{x}_i = x] \tag{1}$$

In this work, we would like to estimate $\mathrm{CACE}(W, Y)$ from data that there exists at least a latent confounder U affecting both W and Y. When there is an IV S and the set of conditioning covariates $\mathbf{Z}$ available in data, $\mathrm{CACE}(W, Y)$ can be calculated by the following formula as in [3,19]:

$$\mathrm{CACE}(W, Y) = \frac{\mathbb{E}(Y|W = 1, S = 1, \mathbf{Z}) - \mathbb{E}(Y|W = 0, S = 1, \mathbf{Z})}{\mathbb{E}(W|S = 1, \mathbf{Z}) - \mathbb{E}(W|S = 0, \mathbf{Z})} \tag{2}$$

The approach of CIV allows a measured covariate to be a valid IV conditioning on a set of measured variables. The formal definition of the CIV in a DAG (Definition 7.4.1 on Page 248 [30]) is introduced as follows.

Definition 1 (Conditional IV). *Let $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ be a DAG with $\mathbf{V} = \mathbf{X} \cup \mathbf{U} \cup \{W, Y\}$, a variable $Q \in \mathbf{X}$ is a conditional IV w.r.t. $W \rightarrow Y$ if there exists a set of measured variables $\mathbf{Z} \subseteq \mathbf{X}$ such that (i) $Q \not\perp\!\!\!\perp_d W \mid \mathbf{Z}$, (ii) $Q \perp\!\!\!\perp_d Y \mid \mathbf{Z}$ in $\mathcal{G}_{\underline{W}}$, and (iii) $\forall Z \in \mathbf{Z}$, Z is not a descendant of Y.*

Here, $\perp\!\!\!\perp_d$ and $\not\perp\!\!\!\perp_d$ are d-separation and d-connection for reading the conditioning relationships between nodes in a DAG [30]. The manipulated DAG $\mathcal{G}_{\underline{W}}$ in Definition 1 is obtained by deleting the direct edge $W \rightarrow Y$ from the DAG $\mathcal{G}$. Note that Definition 1 is defined on a single CIV Q that can be generalised to a set of CIVs $\mathbf{Q}$ easily.

With the pretreatment variables assumption, there is not a descendant of Y in $\mathbf{X}$, i.e. the condition (iii) of Definition 1 is always held. It means that one needs to check the first two conditions for verifying whether a variable is a CIV or not. Note that discovering a conditioning set $\mathbf{Z}$ from a given DAG is NP-complete [37]. Under the pretreatment assumption, the time complexity of discovering a conditioning set is still NP-complete. Instead of discovering a conditioning set from a given causal DAG, in this work, we will utilise disentangled representation learning to learn the representations of CIVs and the representations of the conditioning set directly from data with latent confounders.

3   The Proposed DVAE.CIV Model

3.1   The Disentangled Representation Learning Scheme for Causal Effect Estimation

In this work, we would like to estimate $\text{CACE}(W,Y)$ from observational data with latent confounders. Note that the causal effect of W on Y is non-identifiable when there exists a latent confounder $U \in \mathbf{U}$ affecting both W and Y, i.e. $W \leftarrow U \rightarrow Y$ in the underlying DAG [6,30]. It is challenging to recover $\text{CACE}(W,Y)$ from data with latent confounders due to the effect of U is not computable. If there is a nominated CIV and its corresponding conditioning set, $\text{CACE}(W,Y)$ can be obtained unbiasedly from data by using an IV-based estimator. However, a CIV and its conditioning set are usually unknown in many real-world applications. Furthermore, if an invalid CIV is used, the wrong result or conclusion may be drawn [9,27].

To estimate the conditional average causal effects and average causal effects from data with latent confounders, we propose and design the DVAE.CIV model to learn three representations $\{\mathbf{S}, \mathbf{C}, \mathbf{F}\}$ as in the scheme of Fig. 1(c). Here $\mathbf{S}$ is the representation of CIVs that only affect W but not Y, $\mathbf{F}$ is the representation of the risk factors that affects Y but not W, and $\mathbf{C}$ is the confounding representation that affecting both W and Y.

Our proposed DVAE.CIV model relies on VAEs: we assume that the measured covariates factorise conditioning on the latent variables, and use an inference model [22] which follows a factorisation of the true posterior [15,26]. Based on our disentanglement setting in Fig. 1(c), we have the following theoretical result for causal effect estimation from data with latent confounders.

Theorem 1. *Let $\mathcal{G} = (\mathbf{X} \cup \mathbf{U} \cup \{W,Y\}, \mathbf{E})$ be a causal DAG, in which $\mathbf{X}$ is a set of pretreatment variables, $\mathbf{U}$ is a set of latent confounders, W and Y are treatment and outcome variables, respectively, and $W \rightarrow Y$ is in $\mathbf{E}$. If we can learn the three representations as per the scheme in Fig. 1(c), then the quantities of $CACE(W,Y)$ can be calculated by using IV-based method.*

Proof. The directed edge $W \rightarrow Y$ in $\mathcal{G}$ is to ensure that W has a causal effect on Y. In the causal DAG in Fig. 1(c), we first show that the set $\mathbf{C} \cup \mathbf{F}$ instrumentalists $\mathbf{S}$ to be a valid CIV. $\mathbf{S}$ is a common cause of W and $\mathbf{C}$, so $S \not\perp_d W$, i.e. the first condition of Definition 1 holds. In the causal DAG $\mathcal{G}$ in Fig. 1(c), $\mathbf{C}$ is a collider and is a common cause of W and Y. That is, conditioning on $\mathbf{C}$, the path $W \leftarrow \mathbf{S} \rightarrow \mathbf{C} \leftarrow \mathbf{F} \rightarrow Y$ is open, but $\mathbf{F}$ is sufficient to block this path. For the path $\mathbf{S} \rightarrow \mathbf{C} \rightarrow Y$, $\mathbf{C}$ blocks it. Furthermore, in the manipulated DAG $\mathcal{G}_{\underline{W}}$, W is a collider such that the empty set blocks the three paths between $\mathbf{S}$ and Y, i.e. $\mathbf{S} \rightarrow W \leftarrow U \rightarrow Y$, $\mathbf{S} \rightarrow W \leftarrow \mathbf{C} \leftarrow \mathbf{F} \rightarrow Y$ and $\mathbf{S} \rightarrow W \leftarrow \mathbf{C} \rightarrow Y$. Hence, the set $\mathbf{C} \cup \mathbf{F}$ blocks all paths between $\mathbf{S}$ and Y in $\mathcal{G}_{\underline{W}}$, i.e. the second condition of Definition 1 holds. Finally, $\mathbf{C} \cup \mathbf{F}$ does not contains a descendant of Y due to the pretreatment variables assumption. Thus, the set $\mathbf{C} \cup \mathbf{F}$ instrumentalists $\mathbf{S}$. As in Eq.(2), the IV-based estimators, such as DeepIV [14], can be applied to remove the effect of $\mathbf{U}$ by inputting the CIV representation $\mathbf{S}$ and the representations of its conditioning set $\mathbf{C} \cup \mathbf{F}$. Therefore, the quantities of CACE(W, Y) can be obtained by using the CIV $\mathbf{S}$ and its conditioning set $\mathbf{C} \cup \mathbf{F}$ in an IV-based estimator.

Theorem 1 ensures that a family of data-driven methods can be applied for causal effect estimation from data with latent confounders.

3.2   Learning the Three Representations

Based on Theorem 1, we have known that the set $\{\mathbf{C}, \mathbf{F}\}$ instrumentalists $\mathbf{S}$. In this section, we present our proposed DVAE.CIV model for obtaining the three latent representations from data by using the VAE technique [22], and the architecture of DVAE.CIV is presented in Fig. 2. As shown in Fig. 2, the DVAE.CIV model is to learn and disentangle the latent representation $\mathbf{\Phi}$ of $\mathbf{X}$ into two disjoint sets $\{\mathbf{S}, \mathbf{F}\}$ by using disentangled variational autoencoder [15,38], and generate the representation $\mathbf{C}$ conditioning on $\mathbf{X}$ by jointing the Conditional Variational AutoEncoder (CVAE) network [32].

The DVAE.CIV model is designed to learn three representations shown in Fig. 1(c) by utilising the inference model and generative model to approximate the posterior distribution $p(\mathbf{X}|\mathbf{S}, \mathbf{C}, \mathbf{F})$. The inference model comprises three independent encoders $q(\mathbf{S}|\mathbf{X})$, $q(\mathbf{C}|\mathbf{X})$, and $q(\mathbf{F}|\mathbf{X})$, which are treated as variational posteriors over the three latent representations. The generative model utilises the three latent representations with a decoder model $p(\mathbf{X}|\mathbf{S}, \mathbf{C}, \mathbf{F})$ to reconstruct the measured distribution $\mathbf{X}$.

Following the standard VAE model [22], the prior distributions $p(\mathbf{S})$ and $p(\mathbf{F})$ are drawn from the Gaussian distributions as:

$$p(\mathbf{S}) = \prod_{i=1}^{D_{\mathbf{S}}} \mathcal{N}(S_i|0, 1); \ p(\mathbf{F}) = \prod_{i=1}^{D_{\mathbf{F}}} \mathcal{N}(F_i|0, 1). \tag{3}$$
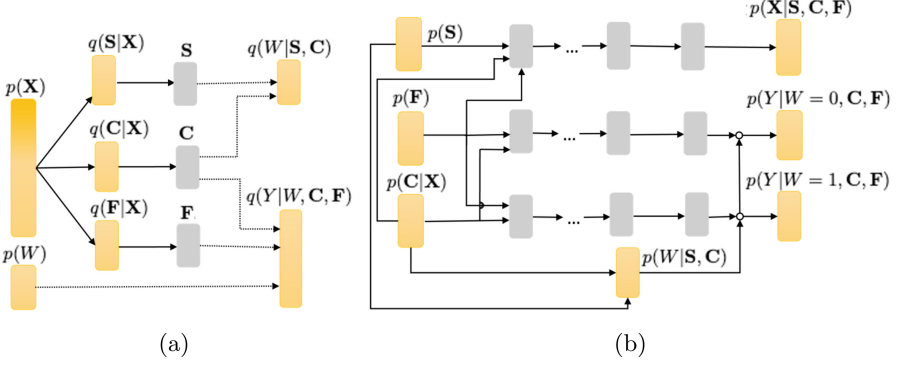
Fig. 2. The architecture of DVAE.CIV model. A yellow box indicates the drawing of samples from the respective distributions, a grey box indicates the parameterised deterministic neural network transitions, and a circle represents switching paths based on the value of W. (Color figure online)

where $D_{\mathbf{S}}$ and $D_{\mathbf{F}}$ are the dimensions of $\mathbf{S}$ and $\mathbf{F}$, respectively. In the inference model, the variational approximations of the posteriors are described as:

$$q(\mathbf{S}|\mathbf{X}) = \prod_{i=1}^{D_{\mathbf{S}}} \mathcal{N}(\mu = \hat{\mu}_{\mathbf{S}_i}, \sigma^2 = \hat{\sigma}_{\mathbf{S}_i}^2); \ q(\mathbf{C}|\mathbf{X}) = \prod_{i=1}^{D_{\mathbf{C}}} \mathcal{N}(\mu = \hat{\mu}_{\mathbf{C}_i}, \sigma^2 = \hat{\sigma}_{\mathbf{C}_i}^2);$$

$$q(\mathbf{F}|\mathbf{X}) = \prod_{i=1}^{D_{\mathbf{F}}} \mathcal{N}(\mu = \hat{\mu}_{\mathbf{F}_i}, \sigma^2 = \hat{\sigma}_{\mathbf{F}_i}^2), \tag{4}$$

where $D_{\mathbf{C}}$ is the dimension of $\mathbf{C}$, and $\hat{\mu}_{\mathbf{S}}, \hat{\mu}_{\mathbf{C}}, \hat{\mu}_{\mathbf{F}}$ and $\hat{\sigma}_{\mathbf{S}}^2, \hat{\sigma}_{\mathbf{C}}^2, \hat{\sigma}_{\mathbf{F}}^2$ are the parameters of means and variances in the Gaussian distributions parameterised by neural networks.

In the generative model, we utilise the Monte Carlo (MC) sampling strategy to sample the distribution $\mathbf{C}$ based on the Conditional Variational AutoEncoder network (CVAE) [32] such that the latent representation of $\mathbf{C}$ is generated from the distribution $\mathbf{X}$:

$$p(\mathbf{C}) \frown p(\mathbf{C}|\mathbf{X}). \tag{5}$$

Furthermore, the generative models for W and $\mathbf{X}$ with the three latent representations are formalised as:

$$p(W|\mathbf{S}, \mathbf{C}) = Bern(\sigma(\psi_1(\mathbf{S}, \mathbf{C}))); \ p(\mathbf{X}|\mathbf{S}, \mathbf{F}) = \prod_{i=1}^{D_{\mathbf{X}}} p(X_i|\mathbf{S}, \mathbf{C}), \tag{6}$$

where $\psi_1(\cdot)$ is a function parameterised by neural networks, $\sigma(\cdot)$ is the logistic function and $Bern(\cdot)$ is the function of Bernoulli distribution.

In our generative model, the latent representation for the outcome Y is based on the data type of Y. For the outcome Y with continuous values, we use a

Gaussian distribution with its mean and variance parameterised by a pair of independent neural networks, i.e. $p(Y|w = 0, \mathbf{C}, \mathbf{F})$ and $p(Y|w = 1, \mathbf{C}, \mathbf{F})$. Thus, the continuous Y is modelled by:

$$
\begin{aligned}
p(Y|W, \mathbf{C}, \mathbf{F}) &= \mathcal{N}(\mu = \hat{\mu}_Y, \sigma^2 = \hat{\sigma}_Y^2), \\
\hat{\mu}_Y &= W \cdot \psi_2(\mathbf{C}, \mathbf{F}) + (1 - W) \cdot \psi_3(\mathbf{C}, \mathbf{F}); \\
\hat{\sigma}_Y^2 &= W \cdot \psi_4(\mathbf{C}, \mathbf{F}) + (1 - W) \cdot \psi_5(\mathbf{C}, \mathbf{F}),
\end{aligned}
\tag{7}
$$

where $\psi_2(\cdot), \psi_3(\cdot), \psi_4(\cdot)$ and $\psi_5(\cdot)$ are neural networks parameterised by their own parameters.

For the outcome Y with binary values, a Bernoulli distribution function based on neural networks is employed to model it and described as:

$$
p(Y|W, \mathbf{C}, \mathbf{F}) = Bern(\sigma(\psi_6(W, \mathbf{C}, \mathbf{F}))),
\tag{8}
$$

where $\psi_6(\cdot)$ is the same with the function ψ_1. These parameters of neural networks can be approximated by maximising the Evidence lower bound (ELBO) $\mathcal{L}_{ELBO}$:

$$
\begin{aligned}
\mathcal{L}_{ELBO}(\mathbf{X}, W, Y) = \ &\mathbb{E}_q[\log p(\mathbf{X}|\mathbf{S}, \mathbf{C}, \mathbf{F})] - D_{KL}[q(\mathbf{S}|\mathbf{X})||p(\mathbf{S})] \\
&- D_{KL}[q(\mathbf{C}|\mathbf{X})||p(\mathbf{C}|\mathbf{X})] - D_{KL}[q(\mathbf{F}|\mathbf{X})||p(\mathbf{F})],
\end{aligned}
\tag{9}
$$

where the decoder $p(\mathbf{C}|\mathbf{X})$ is to ensure that the latent representation $\mathbf{C}$ captures as much information of $\mathbf{X}$ as possible.

To ensure that the treatment W can be recovered from the latent representations $\mathbf{S}$ and $\mathbf{C}$, and the outcome Y can be recovered from the latent representations $\mathbf{C}$ and $\mathbf{F}$, two auxiliary predictors are added and the objective function of DVAE.CIV can be formalised as:

$$
\begin{aligned}
\mathcal{L}_{DVAE.CIV} = \ &-\mathcal{L}_{ELBO}(\mathbf{X}, W, Y) + \alpha\mathbb{E}_q[\log q(W|\mathbf{S}, \mathbf{C})] \\
&+ \beta\mathbb{E}_q[\log q(Y|W, \mathbf{C}, \mathbf{F})],
\end{aligned}
\tag{10}
$$

where α and β are the weights for the auxiliary predictors.

After training the DVAE.CIV model, we get the CIV representation $\mathbf{S}$ and the conditioning set representations $\{\mathbf{C}, \mathbf{F}\}$ based on Theorem 1. For estimating conditional causal effects, we employ an IV-based prediction, DeepIV [14], to implement this part, i.e. we feed $\mathbf{S}$ and $\{\mathbf{C}, \mathbf{F}\}$ into the DeepIV method for conditional causal effect estimation.

4   Experiments

In this section, we evaluate the performance of the proposed DVAE.IV model by applying it to a set of synthetic datasets and three real-world datasets for $CACE(W, Y)$ and average causal effect (ACE) estimation. The three real-world datasets include SchoolingReturns [7], Cattaneo [8] and RHC [11] that are usually utilised in evaluating the methods of causal effect estimation from observational data. Details of the implementation of DVAE.CIV and the appendix are provided in the GitHub[1].

---

[1] https://github.com/IRON13/DVAE.CIV.

4.1   Experimental Setup

We compare the DVAE.CIV against the famous estimators in conditional causal effect estimation that are widely utilised in causal inference from observational data. Note that the ACE can be obtained by averaging the CACE(W, Y) of all individuals. These compared causal effect estimators are introduced in the following.

Compared Causal Effect Estimators. We compare our proposed DVAE.CIV with two Variational AutoEncoder based (VAE-based) causal effect estimators, three tree-based causal effect estimators, two machine learning based (ML-based) causal effect estimators, and three IV-based causal effect estimators. The two VAE-based causal effect estimators are Causal Effect Variational AutoEncoder (CEVAE) [26] and Treatment Effect estimation by Disentangled Variational AutoEncoder (TEDVAE) [38]). The three tree-based causal effect estimators are the standard Bayesian Additive Regression Trees (BART) [17], causal random forest (CF) [35] and causal random forest for IV regression (CFIVR) [4]. Note that CFIVR also belongs to IV-based estimators. The two ML-based causal effect estimators are double machine learning (DML) [10] and doubly robust learning (DRL) [12]. The three IV-based causal effect estimators are DeepIV [14], orthogonal instrumental variable (OrthIV) [33] and double machine learning based IV (DMLIV) [10].

Remarks. The five estimators TEDVAE, BART, CF, DML and DRL rely on the assumption of unconfoundedness [19] (i.e. no latent confounders in data), so the five estimators cannot deal with the case with the data with latent confounders. CEVAE can deal with latent confounders, but it requires that all measured variables are proxy variables of the latent confounders, while our DVAE.CIV model does not have the restriction. The IV-based estimators CFIVR, DeepIV, OrthIV and DMLIV require a known IV that is nominated based on domain knowledge, but the nominated IV usually is not a valid IV and thus may result in a wrong conclusion as argued in Introduction.

Implementation Details. We use *Python* and the libraries including *pytorch* [29], *pyro* [5] and *econml* to implement DVAE.CIV. In our experiments, the dimension of latent representations is set as $|S| = 1$, $|C| = 5$ and $|F| = 5$, respectively. The implementation of CEVAE is based on the *Python* library *pyro* [5] and the code of TEDVAE is from the authors' GitHub[2]. For BART, we use the implementation in the *R* package *bartCause* [17]. For CF and CFIVR, we use the implementations in the *R* functions *causal_forest* and *instrumental_forest* in the *R* package *grf* [4], respectively. The implementations of DML, DRL, DeepIV, OrthIV and DMLIV are from the *Python* package *encoml*.

---

[2] https://github.com/WeijiaZhang24/TEDVAE.

Evaluation Metrics. For performance evaluation, two commonly used metrics are employed in our experiments. For the synthetic datasets, we use absolute error of average causal effect [17], i.e. $\varepsilon_{ACE} = |ACE - A\hat{C}E|$ where ACE is the true causal effect and $A\hat{C}E$ is the estimated causal effect, and Precision of the Estimation of Heterogeneous Effect (PEHE, it is used to evaluate the CACE estimations.) [17,26] $\sqrt{\varepsilon_{PEHE}} = \sqrt{\mathbf{E}(((y_1 - y_0) - (\hat{y}_1 - \hat{y}_0))^2)}$ where y_1, y_0 are the true outcomes and $\hat{y}_1, \hat{y}_0$ are the predicted outcomes, to assess the performance of all methods in terms of the causal effect estimation. Lower values of both metrics indicate better performance. For multiple replications, we present the mean with standard deviation. For the three real-world datasets, we use the reference causal effect in the literature as the baseline to evaluate the performance of all estimators since there is no ground truth causal effect available.

4.2   Simulation Study

It is challenging to evaluate a causal effect estimation method with real-world data since there is no ground truth in the real-world data. In this section, we design simulation studies to evaluate the performance of our proposed DVAE.CIV method in the case that there exists a latent confounder U affecting both W and Y, and there exists a CIV and its conditioning set in the synthetic datasets.

We use a causal DAG $\mathcal{G}$ provided in the appendix to generate synthetic datasets with a range of sample sizes: 2k, 6k, 10k, and 20k. In the causal DAG $\mathcal{G}$, $\mathbf{X} = \{S, X_1, X_2, X_3, X_4, X_5\}$ is the set of measured covariates and $\mathbf{U} = \{U, U_1, U_2, U_3, U_4\}$ is the set of latent confounders in which U affects both W and Y. Note that S is a CIV conditioning on the set $\{X_1, X_2\}$ for all synthetic datasets. Moreover, the data generation process allows the synthetic datasets to have the true individual causal effect. We provide the details of the synthetic data generating process in the appendix. In our experiments, the IV-based estimators OrthIV, DMLIV, DeepIV and CFIVR utilise the true CIV S and the conditioning set $\{X_1, X_2\}$ as input for causal effect estimation.

To provide a reliable assessment, we repeatedly generate 30 synthetic datasets for each sample size setting and utilize the aforementioned metrics to evaluate the performance of the DVAE.CIV against the compared estimators with respect to the task of ACE estimation and CACE estimation from data with latent confounders. For each dataset, we randomly take 70% of samples for training and 30% for testing. The results of all estimators with respect to the ACE estimations and CACE estimations measured by the metrics ε_{ACE} and $\sqrt{\varepsilon_{PEHE}}$ in the out-of-sample set are provided in Tables 1 and 2, respectively. The out-of-sample set is on testing samples, and the within-sample set is on training samples. The results of the within-sample set are provided in the appendix.

Results. By analysing the experiment results in Table 1, we have the following observations: (1) the ML-based and VAE-based estimators, DML, DRL, CEVAE and TEDVAE have the largest ε_{ACE} because the confounding bias caused by confounders and the latent confounder U is not adjusted at all. (2) the tree-based

Table 1. The out-of-sample absolute error ε_{ACE} (mean $\pm$ std) over 30 synthetic datasets. The best results are highlighted in boldface and the runner-up results are underlined. DVAE.CIV is the runner-up on all synthetic datasets, and it relies on the least domain knowledge among all estimators compared since it learns and disentangles the representations of CIV and its conditioning set from data directly.

Samples		2k	6k	10k	20k
Estimators		ε_{ACE}	ε_{ACE}	ε_{ACE}	ε_{ACE}
ML-based	DML	5.507 ± 0.387	5.624 ± 0.182	5.619 ± 0.122	5.633 ± 0.096
	DRL	5.746 ± 0.404	5.833 ± 0.186	5.825 ± 0.156	5.860 ± 0.106
tree-based	BART	3.586 ± 0.179	3.596 ± 0.090	3.613 ± 0.065	3.622 ± 0.060
	CF	3.226 ± 0.342	3.246 ± 0.141	3.274 ± 0.127	3.312 ± 0.074
VAE-based	CEVAE	$5.595 \ \pm 0.455$	5.652 ± 0.183	5.631 ± 0.179	5.726 ± 0.123
	TEDVAE	5.615 ± 0.455	5.655 ± 0.181	5.634 ± 0.172	5.696 ± 0.112
IV-based	OrthIV	2.212 ± 1.260	1.952 ± 0.585	1.792 ± 0.607	1.974 ± 0.419
	DMLIV	2.170 ± 1.189	1.888 ± 0.572	1.790 ± 0.626	1.971 ± 0.432
	DeepIV	$\mathbf{0.352 \pm 0.180}$	0.632 ± 0.245	0.726 ± 0.315	0.757 ± 0.354
	CFIVR	1.228 ± 0.949	$\mathbf{0.504 \pm 0.369}$	$\mathbf{0.543 \pm 0.474}$	$\mathbf{0.415 \pm 0.307}$
DVAE.CIV		$\underline{0.577 \pm 0.117}$	$\underline{0.577 \pm 0.064}$	$\underline{0.561 \pm 0.075}$	$\underline{0.512 \pm 0.091}$

estimators, BART and CF have the second largest ε_{ACE} as they fail to deal with the confounding bias caused by the latent confounder U. (3) the IV-based estimators including DVAE.CIV significantly outperform the other estimators including DML, DRL, BART, CF, CEVAE and TEDVAE. (4) DVAE.CIV is the second best performer on all synthetic datasets and its performance is comparable with CFIVR and DeepIV. (5) as the sample size increases, the standard deviation of most estimators including DVAE.CIV decreases significantly. It's worth mentioning that DVAE.CIV requires the least domain knowledge among all estimators since it only relies on the assumption that there exists a CIV and the conditioning set (maybe an empty set). This is very important in practice, as in many real-world applications, there is rarely sufficient prior knowledge for nominating a valid IV.

From the results in Table 2, we can conclude that (1) the ML-based, tree-based, and VAE-based estimators have the worst performance with respect to conditional causal effect estimations. (2) Among the IV-based estimators, DeepIV achieves the best performance on the first two groups of synthetic datasets and the second-best performance on the other datasets, and CFIVR obtains the best performance on the last four groups of synthetic datasets and the second-best performance on the first two groups of synthetic datasets. (3) DVAE.CIV obtains the second-best performance on all synthetic datasets. (4) The standard deviation of DVAE.CIV is the smallest on all datasets, and as the sample size increases, the standard deviation of DVAE.CIV reduces significantly. These conclusions demonstrate that DVAE.CIV can learn and disentangle the

Table 2. The out-of-sample $\sqrt{\varepsilon_{PEHE}}$ (mean $\pm$ std) over 30 synthetic datasets. The lowest $\sqrt{\varepsilon_{PEHE}}$ are highlighted in boldface and the runner-up results are underlined. DVAE.CIV is in the runner-up results on the first two groups of synthetic datasets and achieves the third smallest $\sqrt{\varepsilon_{PEHE}}$ on the last four groups of synthetic datasets. It's worth mentioning that DVAE.CIV obtains the lowest standard deviation on all synthetic datasets.

Samples		2k	6k	10k	20k
Estimators		$\sqrt{\varepsilon_{PEHE}}$	$\sqrt{\varepsilon_{PEHE}}$	$\sqrt{\varepsilon_{PEHE}}$	$\sqrt{\varepsilon_{PEHE}}$
ML-based	DML	5.484 ± 0.382	5.584 ± 0.167	5.580 ± 0.128	5.609 ± 0.105
	DRL	5.701 ± 0.408	5.773 ± 0.179	5.767 ± 0.156	5.815 ± 0.112
tree-based	BART	4.791 ± 0.205	4.790 ± 0.083	4.789 ± 0.072	4.790 ± 0.060
	CF	3.483 ± 0.319	3.500 ± 0.134	3.523 ± 0.120	3.554 ± 0.070
VAE-based	CEVAE	6.093 ± 0.396	6.138 ± 0.175	6.107 ± 0.160	6.192 ± 0.112
	TEDVAE	6.111 ± 0.392	6.138 ± 0.177	6.110 ± 0.158	6.167 ± 0.103
IV-based	OrthIV	3.070 ± 0.718	2.798 ± 0.299	2.734 ± 0.256	2.795 ± 0.218
	DMLIV	3.027 ± 0.682	2.767 ± 0.278	$2.736 \pm .0.268$	2.794 ± 0.221
	DeepIV	$\mathbf{2.396 \pm 0.054}$	$\underline{2.412 \pm 0.042}$	$\mathbf{2.418 \pm 0.060}$	$\underline{2.425 \pm 0.065}$
	CFIVR	3.016 ± 0.658	$\mathbf{2.421 \pm 0.235}$	$\underline{2.423 \pm 0.351}$	$\mathbf{2.203 \pm 0.145}$
DVAE.CIV		$\underline{2.448 \pm 0.044}$	2.460 ± 0.037	2.452 ± 0.024	2.442 ± 0.025

representations of the CIV and its conditioning set for CACE estimation from data with latent confounders.

In conclusion, DVAE.CIV achieves competitive performance compared to state-of-the-art causal effect estimators while requiring the least prior knowledge in ACE and CACE estimations from observational data with latent confounders.

4.3   Experiments on Three Real-World Datasets

We selected three real-world datasets with their empirical causal effect values available and commonly used in the literature to assess the performance of DVAE.CIV in ACE estimations. We did not conduct experiments on CACE estimation on the three datasets since there were no ground truth or empirical estimates of CACEs available for these datasets. The three real-world datasets are SchoolingReturns [7], Cattaneo [8], and RHC [11]. These datasets are widely utilized in the evaluation of either IV estimators or data-driven causal effect estimators [13]. Note that SchoolingReturns has a nominated CIV, and the last two datasets do not have a nominated IV for causal effect estimation. Thus, we only compared the DVAE.CIV model with all the aforementioned estimators on SchoolingReturns and the ML-based, tree-based, and VAE-based estimators on both Cattaneo and RHC datasets.

SchoolingReturns. The dataset is from the national longitudinal survey of youth (NLSY), a well-known dataset of US young employees, aged range from 24 to

Table 3. Estimated ACEs by all methods on the three real-world datasets. We highlight the estimated causal effects within the empirical interval on SchoolingReturns and Cattaneo. We use '-' to indicate that an IV-based estimator does not work on Cattaneo and RHC since there is not a nominated IV. Note that all estimators on RHC obtain a consistent result.

Samples		SchoolingReturns	Cattaneo	RHC
ML-based	DML	−0.0227	−150.21	**0.0244**
	DRL	−0.0154	−164.32	**0.0447**
tree-based	BART	−0.0384	−172.53	**0.0381**
	CF	**0.1400**	**−232.33**	0.0278
VAE-based	CEVAE	0.02617	**−221.23**	0.0322
	TEDVAE	0.0029	**−228.65**	0.0293
IV-based	OrthIV	1.3180	–	–
	DMLIV	1.2806	–	–
	DeepIV	0.0328	–	–
	CFIVR	1.1510	–	–
DVAE.CIV		**0.1855**	**−224.79**	**0.0414**

34 [7]. The dataset has 3,010 samples and 19 variables [7]. The variable of the education of employees is the treatment variable, and the variable of the raw wages in 1976 (in cents per hour) is the outcome variable. The dataset was collected to study the causal effect of education on earnings. Note that the variable of geographical proximity to a college, i.e. *nearcollege* is nominated to be an IV by Card [7]. The empirical estimate $ACE(W, Y) = 0.1329$ with 95% confidence interval $(0.0484, 0.2175)$ is from [34] and used as the reference value.

Cattaneo. The dataset has the birth weights of 4,642 singleton births with 20 variables ([8]) that were collected from Pennsylvania, USA for the study of the average of maternal smoking status during pregnancy (W) on a baby's birth weight $(Y$, in grams). The dataset contains several covariates: mother's age, mother's marital status, an indicator for the previous infant where the newborn died, mother's race, mother's education, father's education, number of prenatal care visits, months since last birth, an indicator of firstborn infant and indicator of alcohol consumption during pregnancy. The authors [8] found a strong negative effect of maternal smoking on the weights of babies, i.e., about 200g to 250g lighter for a baby with a mother smoking during pregnancy.

Right Heart Catheterization (RHC). RHC is a real-world dataset obtained from an observational study regarding a diagnostic procedure for the management of critically ill patients [11]. The RHC dataset can be downloaded from the **R** package *Hmisc*[3]. The dataset contains 2,707 samples with 72 covariates [11,25].

---

[3] https://CRAN.R-project.org/package=Hmisc.

RHC was for investigating the adult patients who participated in the Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT). The treatment variable W is whether a patient received an RHC within 24 h of admission, and the outcome variable Y is whether a patient died at any time up to 180 d after admission. Note that the empirical conclusion is that applying RHC leads to higher mortality within 180 d than not applying RHC [11].

Results. All results on the three real-world datasets are reported in Table 3. From Table 3, we make the following observations: (1) the estimated causal effects by DVAE.CIV and CF on SchoolingReturns and Cattaneo fall within the empirical intervals, while DML, DRL, and BART provide an opposite estimate to the empirical value on SchoolingReturns; (2) as there is no nominated IV on Cattaneo and RHC, the estimators OrthIV, DMLIV, DeepIV, and CFIVR do not work on both datasets; (3) all estimators, including DVAE.CIV, obtain a consistent estimation on the RHC data, and they reach the same conclusion as the empirical conclusion [11]. These observations further confirm that DVAE.CIV is capable of removing the bias between W and Y in real-world datasets.

In conclusion, our simulation studies show the high performance of DVAE.CIV in ACE and CACE estimations from data with latent confounders, and our experiments on three real-world datasets further confirm the capability of DVAE.CIV in ACE estimation from observational data.

Limitations. The performance of DVAE.CIV relies on the assumptions made in this work and the assumptions on the VAE model. Note that the identification of the VAE model [21] is an important issue for our proposed DVAE.CIV model. When some of the assumptions or the VAE identification do not hold, DVAE.CIV may obtain an inconsistent conclusion. To obtain a consistent conclusion, it would be better to conduct a sensitivity analysis [19,30] together with DVAE.CIV to achieve a reliable conclusion in real-world applications.

5   Conclusion

It is a crucial challenge to deal with the bias caused by latent confounders in conditional causal effect estimations from observational data. IV-based methods allow us to remove such confounding bias in an effective way, but it relies on a nominated IV/CIV based on domain knowledge. In this paper, we propose an efficient approach, DVAE.CIV for conditional causal effect estimations from observational data with latent confounders. The DVAE.CIV utilizes the advantages of deep generative models for learning the representations of a CIV and its conditioning set from data with latent confounders. We theoretically show the soundness of the DVAE.CIV model. The effectiveness and potential of the DVAE.CIV are demonstrated by extensive experiments. In simulation studies, DVAE.CIV achieves competitive performance against state-of-the-art estimators that require extra prior knowledge in ACE and CACE estimation from data with

latent confounders. The experimental results on three real-world datasets show the superiority of the DVAE.CIV model on ACE estimation over the existing estimators.

References

1. Abadie, A., Imbens, G.W.: Large sample properties of matching estimators for average treatment effects. Econometrica **74**(1), 235–267 (2006)
2. Angrist, J.D., Imbens, G.W.: Two-stage least squares estimation of average causal effects in models with variable treatment intensity. J. Am. Stat. Assoc. **90**(430), 431–442 (1995)
3. Angrist, J.D., Imbens, G.W., Rubin, D.B.: Identification of causal effects using instrumental variables. J. Am. Stat. Assoc. **91**(434), 444–455 (1996)
4. Athey, S., Tibshirani, J., Wager, S.: Generalized random forests. Ann. Stat. **47**(2), 1148–1178 (2019)
5. Bingham, E., Chen, J.P., et al.: Pyro: deep universal probabilistic programming. J. Mach. Learn. Res. **20**(1), 973–978 (2019)
6. Brito, C., Pearl, J.: Generalized instrumental variables. In: Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence, pp. 85–93 (2002)
7. Card, D.: Using geographic variation in college proximity to estimate the return to schooling (1993)
8. Cattaneo, M.D.: Efficient semiparametric estimation of multi-valued treatment effects under ignorability. J. Econometrics **155**(2), 138–154 (2010)
9. Cheng, D., Li, J., Liu, L., Liu, J., Le, T.D.: Data-driven causal effect estimation based on graphical causal modelling: A survey. arXiv preprint arXiv:2208.09590 (2022)
10. Chernozhukov, V., et al.: Double/debiased machine learning for treatment and structural parameters. Econometrics J. **21**(1), C1–C68 (2018)
11. Connors, A.F., Speroff, T., et al.: The effectiveness of right heart catheterization in the initial care of critically III patients. J. Am. Med. Assoc. **276**(11), 889–897 (1996)
12. Foster, D.J., Syrgkanis, V.: Orthogonal statistical learning. arXiv preprint arXiv:1901.09036 (2019)
13. Guo, R., Cheng, L., Li, J., Hahn, P.R., Liu, H.: A survey of learning causality with data: problems and methods. ACM Comput. Surv. (CSUR) **53**(4), 1–37 (2020)
14. Hartford, J., Lewis, G., et al.: Deep IV: A flexible approach for counterfactual prediction. In: International Conference on Machine Learning, pp. 1414–1423 (2017)
15. Hassanpour, N., Greiner, R.: Learning disentangled representations for counterfactual regression. In: International Conference on Learning Representations, pp. 1–11 (2019)
16. Hernán, M.A., Robins, J.M.: Instruments for causal inference: an epidemiologist's dream? Epidemiology **17**(4), 360–372 (2006)
17. Hill, J.L.: Bayesian nonparametric modeling for causal inference. J. Comput. Graph. Stat. **20**(1), 217–240 (2011)
18. Imbens, G.W.: Instrumental variables: an econometrician's perspective. Stat. Sci. **29**(3), 323–358 (2014)

19. Imbens, G.W., Rubin, D.B.: Causal Inference in Statistics, Social, and Biomedical Sciences. Cambridge University Press, Cambridge (2015)
20. Kang, H., Zhang, A., Cai, T.T., Small, D.S.: Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. J. Am. Stat. Assoc. **111**(513), 132–144 (2016)
21. Khemakhem, I., Kingma, D., Monti, R., Hyvarinen, A.: Variational autoencoders and nonlinear ICA: a unifying framework. In: International Conference on Artificial Intelligence and Statistics, pp. 2207–2217. PMLR (2020)
22. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In: International Conference on Learning Representations (2014)
23. Kuang, Z., Sala, F., et al.: Ivy: instrumental variable synthesis for causal inference. In: International Conference on Artificial Intelligence and Statistics, pp. 398–410 (2020)
24. LaLonde, R.J.: Evaluating the econometric evaluations of training programs with experimental data. Am. Econ. Rev. **76**(4), 604–620 (1986)
25. Loh, W.W., Vansteelandt, S.: Confounder selection strategies targeting stable treatment effect estimators. Stat. Med. **40**(3), 607–630 (2021)
26. Louizos, C., Shalit, U., Mooij, J.M., Sontag, D., Zemel, R., Welling, M.: Causal effect inference with deep latent-variable models. In: Advances in Neural Information Processing Systems, pp. 6446–6456 (2017)
27. Martens, E.P., Pestman, W.R., de Boer, A., Belitser, S.V., Klungel, O.H.: Instrumental variables: application and limitations. Epidemiology **17**(3), pp. 260–267 (2006)
28. Martinussen, T., Nørbo Sørensen, D., Vansteelandt, S.: Instrumental variables estimation under a structural cox model. Biostatistics **20**(1), 65–79 (2019)
29. Paszke, A., Gross, S., et al.: Pytorch: an imperative style, high-performance deep learning library. In: International Conference on Neural Information Processing Systems, pp. 8026–8037 (2019)
30. Pearl, J.: Causality. Cambridge University Press, Cambridge (2009)
31. Rosenbaum, P.R., Rubin, D.B.: The central role of the propensity score in observational studies for causal effects. Biometrika **70**(1), 41–55 (1983)
32. Sohn, K., Yan, X., Lee, H.: Learning structured output representation using deep conditional generative models. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. vol. 2, pp. 3483–3491 (2015)
33. Syrgkanis, V., Lei, V., et al.: Machine learning estimation of heterogeneous treatment effects with instruments. In: International Conference on Neural Information Processing Systems, pp. 15193–15202 (2019)
34. Verbeek, M.: A Guide to Modern Econometrics. Wiley, Hoboken (2008)
35. Wager, S., Athey, S.: Estimation and inference of heterogeneous treatment effects using random forests. J. Am. Stat. Assoc. **113**(523), 1228–1242 (2018)
36. Yuan, J., Wu, A., et al.: Auto IV: counterfactual prediction via automatic instrumental variable decomposition. ACM Trans. Knowl. Discov. Data **16**(4), 1–20 (2022)
37. Van der Zander, B., Liśkiewicz, M., Textor, J.: Efficiently finding conditional instruments for causal inference, pp. 3243–3249 (2015)
38. Zhang, W., Liu, L., Li, J.: Treatment effect estimation with disentangled latent factors. In: The AAAI Conference on Artificial Intelligence, pp. 10923–10930 (2021)