Causal Risk Minimization for High-Dimensional Treatments

Nikita Dhawan [1 2] **Arnav Paruthi** [1] **Andrew Kim** [1 2] **Lovedeep Gondara** [3] **Jekaterina Novikova** [3]
Chris J. Maddison [1 2]

Abstract

Predicting the effect of interventions with many possible variations, *e.g.*, therapeutic content that affects mental health outcomes or an earnings call transcript that drives movement in share price, is useful across several domains. However, classical causal estimators tend to assume that all possible interventions are observed, which is infeasible when interventions vary widely, for instance, in the space of all text strings. We adapt a well-known approach of recasting causal inference as a learning problem, to address high-dimensional treatment spaces. Specifically, under standard assumptions like no unobserved confounding, we show that causal error decomposes into a series of moment-balancing errors of increasing order, and design objectives that directly improve causal estimation. We also show how to project the effect of a high-dimensional treatment onto lower-dimensional treatment attributes, which allows a single model to answer several causal questions without additional attribute-specific training. We empirically evaluate our estimators in settings with high-dimensional continuous, discrete, and text treatments, the last of which used a semi-synthetic dataset of Amazon Reviews. Our experiments demonstrate the benefit of higher-order balance error optimization and competitive performance of projected causal estimates with attribute-specific estimators. Corresponding code is available here.

1. Introduction

Effective decision-making in real-world high-stakes domains often relies on predicting the effects of a combination of many different choices. For instance, the success of a financial advisor's client call may depend on the products recommended, the order in which information is presented, the style of language used, and how these interact. Understanding the effect of all of these elements of a call on a single outcome, *e.g.*, customer satisfaction, is a causal estimation problem with a large space of possible interventions (or treatments), which in this case is the space of all strings. Other examples include a counselor's response to a patient in distress (Althoff et al., 2016; Ewbank et al., 2020), the phrasing of a public health message (Milkman et al., 2021; Dai et al., 2021), word choices in advertisements (Pryzant et al., 2018), and the framing of an earnings disclosure (Hassan, 2020).

In this work, we focus on estimating the *average potential outcome (APO)* of a high-dimensional treatment, which is defined as the average outcome that would be obtained under that treatment in a given population in the Neyman–Rubin model (Holland, 1986; Hirano et al., 2003). Text strings as treatments is the key motivating setting, since understanding the effect of a whole text intervention can enable many types of decision-making: selecting optimal wording combinations, ranking candidate messages, or identifying linguistic strategies to achieve an outcome. Specifically, our setting is challenging due to the dimensionality of the treatment because each observed data point may have received a unique treatment and test-time treatments may be unobserved during training.

The traditional causal estimation approaches struggle when faced with high-dimensional treatments. Randomized controlled trials are too expensive or infeasible to conduct while covering the entire space of possible treatments. Naïve applications of classical observational estimators to this setting also face several challenges. Inverse propensity score weighting (IPW) (Horvitz & Thompson, 1952; Rosenbaum, 1987) computes an empirical sum that is restricted to units that received the exact treatment of interest, and is undefined for any string not observed in the training data. Outcome imputation (Rubin, 1977) can sidestep this by learning a model of the conditional outcome given treatment and confounders and marginalizing over the confounders, but this requires inference-time access to individual-level confounder values, which may be impractical when those variables are sensitive (*e.g.*, income or health markers) and prohibitively expensive when the confounders themselves are high-dimensional

[1]University of Toronto [2]Vector Institute [3]Vanguard. Correspondence to: Nikita Dhawan <nikita@cs.toronto.edu>.

ICML 2026 Workshop on Hypothesis Testing, Seoul, South Korea, 2026. Copyright 2026 by the author(s).

(*e.g.*, textual confounders).

The approach we adapt is to recast causal estimation as a learning problem so that generalization handles previously unseen treatments. It is known that causal estimation can be recast as weighted risk minimization (Swaminathan & Joachims, 2015; Hassanpour & Greiner, 2019; Jung et al., 2020), a general approach we call *Causal Risk Minimization (CRM)* throughout, but works in the CRM literature tend to focus on binary treatments, which obscures some of the challenges with high-dimensional treatments.

Our key contribution is to address two of the challenges faced by causal risk minimization with high-dimensional treatments. First, APO estimation via risk minimization requires targets that must themselves be estimated from data. In the case of IPW, the targets are weighted outcomes, such that APO estimation quality depends directly on the correctness of the learned weights. Empirical or trained models of these weights often run the risk of producing extreme weights which in turn harm APO training and prediction. We address this by showing that the APO estimation error decomposes into a series of *moment-balancing errors* — errors in balancing the moments of different orders of the confounder distribution across treatment groups. This decomposition, derived for binary, discrete, and continuous confounders, motivates learning importance weights with explicit balance regularization up to a chosen order. While prior works have used low-order balance regularization to improve propensity score training (Imai & Ratkovic, 2014), our error decomposition provides a principled strategy to choose the order upto which balancing confounder moments improves APO estimation itself.

Second, practical use-cases of an APO estimator may include asking lower-dimensional causal questions about the effects of specific text attributes on an outcome. Prior work using texts as treatments has largely circumvented high-dimensionality issues by collapsing the text into a low-dimensional, often binary, summary before classical estimation (Fong & Grimmer, 2016; Pryzant et al., 2018, 2020; Zhang et al., 2020). However, this approach may sacrifice valuable information, requires committing to the textual attributes of interest *a priori*, and involves training a separate estimator per attribute. Instead, we demonstrate how to efficiently project a single high-dimensional APO estimator trained over the full (text) treatment space onto lower-dimensional attributes *post hoc*, given access to a model of the relationship between the attributes and text. This allows us to simultaneously answer several lower-dimensional causal questions, at no additional training cost per attribute.

We evaluated our CRM estimators across three settings of increasing complexity: a simple linear setting with continuous treatments, a synthetic dataset with high-dimensional discrete treatments, and a semi-synthetic dataset of real Amazon Reviews (Hou et al., 2024) with 10,000 unique text treatments. For text treatments, we finetuned a pretrained language model on the CRM loss function that is aligned with common pretraining objectives and hence, can harness the priors that these models encode over natural language (Choi et al., 2022; Capstick et al., 2025). Our experiments empirically verify the analytic APO error decomposition, demonstrate generalization to unseen treatments, and tease apart the effect of different experimental parameters on prediction quality. In particular, we found that optimizing higher-order moment-balancing errors consistently improved performance of our weighted CRM estimator, outperforming all other variants, without access to confounders at inference-time. Finally, projections of our Amazon Reviews APO estimator onto attributes like sentiment and review length matched or outperformed bespoke estimators that were trained separately per attribute.

2. Preliminaries

Let $\mathcal{T}$ be the space of possible treatments and let $Y(t)$ for $t \in \mathcal{T}$, X, and T be random variables corresponding to the potential outcomes, the confounders, and the treatment, respectively, all drawn from some joint distribution. We assume $Y(t)$ takes values in either $\mathbb{R}$ or $\{0, 1\}$ and X takes values in $\mathcal{X}$. Our goal is to estimate the *average potential outcome* (APO) under treatment t,

$$g(t) = \mathbb{E}[Y(t)], \tag{1}$$

as defined in the Neyman–Rubin model (Holland, 1986). We assume access to n i.i.d. samples (Y_i, X_i, T_i) from an observational distribution, where we abuse notation and let $Y = Y(T)$ be the random potential outcome realized under the observed treatment T. Since the treatment assignment may depend on X, the naïve conditional mean $\mathbb{E}[Y \mid T = t]$ is in general biased for $g(t)$, as illustrated in Figure 1. Throughout, we assume strong ignorability (*i.e.* no unmeasured confounding): $\{Y(t)\}_{t\in\mathcal{T}} \perp T \mid X$, positivity: $0 < p_{T|X}(t \mid x) < 1$, and the stable unit treatment value assumption (Rubin, 1980), under which the following classical estimators are valid. We provide a complete list of notation in Section A.

Inverse Propensity Score Weighting (IPW). The IPW estimator (Horvitz & Thompson, 1952; Rosenbaum, 1987) uses an estimate $\hat{e}(t, x)$ of the *propensity score* $p_{T|X}(t \mid x)$ to estimate APOs as:

$$\hat{g}_{\text{IPW}}(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i \, \mathbb{I}\Big(T_i = t\Big)}{\hat{e}(T_i, X_i)}. \tag{2}$$

Notice that the indicator $\mathbb{I}(T_i = t)$ restricts the sum above to units that received treatment t. In high-dimensional settings like those with text treatments, we may only ever observe

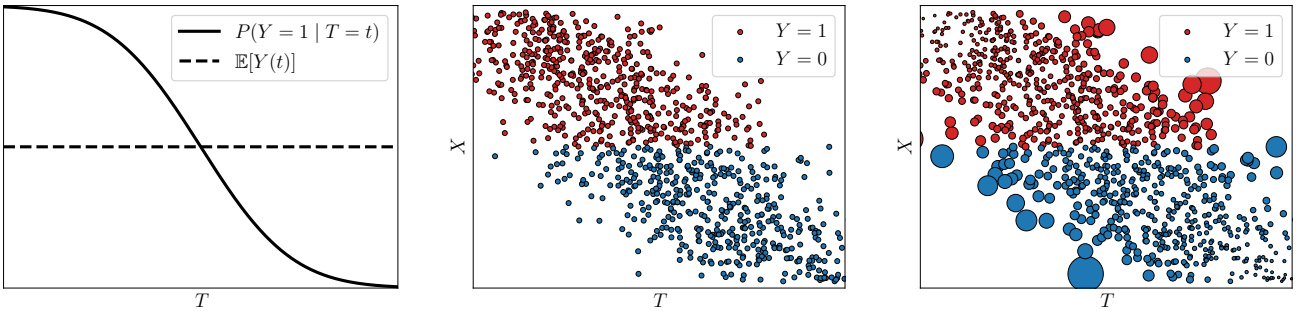

Figure 1. In this toy example with confounding due to X, the conditional mean $P(Y = 1 \mid T = t)$ diverges from the true APO $\mathbb{E}[Y(t)]$, indicating an effect of T when there isn't one (**Left**). Naive ERM using samples (**Middle**) approximates this biased conditional, while weighted risk minimization, (**Right**) with correct weights, visualized by marker sizes, correctly recovers the true APO function.

one datapoint per treatment and for treatments not seen in the training data, this estimator is undefined.

Outcome Imputation (OI). The outcome imputation estimator (Rubin, 1977) learns an estimate $\hat{f}(t, x)$ of the expected conditional outcome $\mathbb{E}[Y \mid T = t, X = x]$ to estimate APOs as:

$$\hat{g}_{\text{OI}}(t) = \frac{1}{n} \sum_{i=1}^{n} \hat{f}(t, X_i). \tag{3}$$

While $\hat{f}$ may generalize to unseen t with a correctly specified model class, the Monte Carlo average over n confounder values for each treatment t can be computationally expensive at inference time. Moreover, when confounders X are sensitive (*e.g.*, income or health status), access to individual-level X_i at inference time may not be feasible for privacy reasons.

These observations motivate the development of estimators that (i) can generalize across treatments by exploiting the structure in $\mathcal{T}$ and the inductive bias of trained models, and (ii) produce APO estimates $\hat{g}(t)$ as a direct function of t alone, without requiring access to confounders at test time.

3. Causal Risk Minimization

Many causal estimation problems reduce to risk minimization problems (Swaminathan & Joachims, 2015; Hassanpour & Greiner, 2019; Jung et al., 2020), approaches that we collectively call Causal Risk Minimization (CRM). Figure 1 demonstrates with a toy example how weighted risk minimization can recover true APOs given correct weights. The CRM approach that we take is to derive a supervised learning problem whose risk minimizer is the APO function. The critical quantity in this problems is a weight function, whose estimation is the focus of one of our contributions. Our second key contribution is a simple Monte Carlo method for computing APO functions of treatments projected to lower dimensions. For text treatments in particular, the cross-entropy loss variant derived below allows us to finetune pretrained language models on an objective that aligns with their next-token-prediction pretraining objective.

3.1. Deriving CRM Targets

IPW-CRM. The IPW estimator in Equation (2) is limited by the requirement that treatment t appear in the observed data. Note that, under strong ignorability and with true propensity scores,

$$\begin{aligned} g(t) &= \mathbb{E}\left[\frac{Y\mathbb{I}\Big(T = t\Big)}{p_{T|X}(T \mid X)}\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{Y\mathbb{I}\Big(T = t\Big)}{p_{T|X}(T \mid X)} \,\middle|\, T\right]\right] \\ &= \int p_T(s)\mathbb{E}\left[\frac{Y\mathbb{I}\Big(T = t\Big)}{p_{T|X}(T \mid X)} \,\middle|\, T = s\right] ds \\ &= \mathbb{E}\left[\frac{Y p_T(t)}{p_{T|X}(t \mid X)} \,\middle|\, T = t\right], \end{aligned} \tag{4}$$

where the first three equalities follow, respectively, from the standard IPW estimator, the law of total expectation, and the definition of expectation with respect to T. The final equality holds because in the integral over $s \in \mathcal{T}$, the indicator variable $\mathbb{I}(T = t)$ evaluates to 1 when $s = t$ and 0 otherwise. Equation (4) motivates fitting a parametric IPW-CRM model $\hat{g}(t)$ with the following objective:

$$\hat{g}_{\text{IPW-CRM}} = \arg\min_{g} \sum_{i=1}^{n} \mathcal{L}\left(g(T_i), \frac{\hat{p}_T(T_i)}{\hat{e}(T_i, X_i)} Y_i\right), \tag{5}$$

where $\hat{e}(T_i, X_i)$ is a learned propensity score, $\hat{p}_T = \sum_{i=1}^{n} \hat{e}(T_i, X_i)$ is the marginal treatment probability estimate, and $\mathcal{L}$ is a suitable loss function. For real Y, we can take $\mathcal{L}$ as squared error, and for binary Y, we take it as cross-entropy loss with targets $[1 - Y_i \cdot \frac{\hat{p}_T(T_i)}{\hat{e}(T_i, X_i)}, Y_i \cdot \frac{\hat{p}_T(T_i)}{\hat{e}(T_i, X_i)}]$.

SW-CRM. The weighted-outcome targets in Equation (5) are computed by first learning a model for propensity scores. However, we may also directly learn a model of the stabilized weights $\hat{w}(T_i, X_i)$ (Robins et al., 2000) to then fit a

Stabilized-Weighted APO estimator (SW-CRM) given by:

$$\hat{g}_{\text{SW-CRM}} = \arg\min_{g} \sum_{i=1}^{n} \mathcal{L}\left(g(T_i), \hat{w}(T_i, X_i)\, Y_i\right), \quad (6)$$

where the loss function $\mathcal{L}$ is chosen similarly to that for IPW-CRM.

OI-CRM. The OI estimator of Equation (3) required fitting a conditional outcome model $\hat{f}$ and then estimating APOs with $\hat{g}(t) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}(t, X_i)$, assuming access to individual-level confounder values and requiring multiple (n) forward passes during inference. Given a trained $\hat{f}$, we can lift this requirement by training an OI-CRM estimator:

$$\hat{g}_{\text{OI-CRM}} = \arg\min_{g} \sum_{i=1}^{n} \mathcal{L}\left(g(T_i), \frac{1}{n}\sum_{i=1}^{n} \hat{f}(T_i, X_i)\right), \quad (7)$$

where the loss function $\mathcal{L}$ is squared error. Note that all estimators learned via Equations (5) to (7) learn a direct mapping from treatments to APOs, do not assume access to confounder values at inference time, and require only a single forward pass to predict the APO for a given treatment.

3.2. APO Error Decomposition

The solutions of Equations (5) to (7) rely crucially on the targets used in the loss function, which are themselves computed using a learned model of propensity scores, stabilized weights, or conditional outcomes, respectively. However, even small errors in calibration of these learned models, a common pitfall of neural networks, can lead to large errors in APO estimation. In the case of propensity scores and stabilized weights, we can decompose the error in APO estimation into a series of balancing error terms. Confounder balancing is a well-known necessary condition for correct propensity scores (Imai & Ratkovic, 2014). The error decomposition below shows the sufficiency of minimizing balance error upto some order for minimizing APO estimation error via risk minimization. We present the decomposition in terms of the stabilized weights $w(t, x)$, while the same analysis holds for propensity scores $e(t, x)$.

The balancing property of stabilized weights states that they balance any function h of the confounders $X \in \mathcal{X}$ across treatments: $\mathbb{E}[w(t, X)\, h(X)] = \mathbb{E}[h(X)]$ for all $t \in \mathcal{T}$. Perfect APO estimation via the optimization in Equation (6) requires the learned $\hat{w}(t, X)$ to satisfy this for $h(X) = f(t, X)$.

$$\begin{aligned}
\hat{g}_{\text{SW-CRM}}(t) &= \mathbb{E}[\hat{w}(T, X)\, Y \mid T = t] \\
&= \mathbb{E}[\mathbb{E}[\hat{w}(T, X) \cdot Y \mid T = t, X] \mid T = t] \\
&= \mathbb{E}[\hat{w}(t, X)\mathbb{E}[Y \mid T = t, X] \mid T = t] \\
&= \mathbb{E}[\hat{w}(t, X)\, f(t, X) \mid T = t].
\end{aligned}$$

Hence,

$$\begin{aligned}
&\hat{g}_{\text{SW-CRM}}(t) = g(t) = \mathbb{E}[f(t, X)] \\
&\iff \mathbb{E}[\hat{w}(t, X)\, f(t, X) \mid T = t] = \mathbb{E}[f(t, X)].
\end{aligned} \quad (8)$$

In other words, the APO estimation error depends directly on how well the weights $\hat{w}(t, X)$ balance $f(t, X)$ for each treatment t. Without access to the true f, existing work attempts to learn weights that balance either arbitrary functions of X (Du et al., 2021) or, more commonly, the first moment of X (Athey et al., 2018). Under the following assumption on the structure of $f(t, \cdot)$, we can relate APO estimation error to moment-balancing errors, such that minimizing balance errors directly minimizes APO error.

Assumption 1. *Either the conditional outcome function $f(t, \cdot)$ is a polynomial of degree at most K for all $t \in \mathcal{T}$, or confounders X are supported on a finite set $\{0, 1, \ldots, K\}$.*

Under Assumption 1, we can write $f(t, X) = \sum_{k=0}^{K} c_k(t)\, X^k)$ for constants $c_k(t)$, and Equation (8) gives:

$$\hat{g}(t) - g(t) = \sum_{k=0}^{K} c_k(t) \cdot \varepsilon_k(t), \quad (9)$$

where $\varepsilon_k(t) = \mathbb{E}\left[\hat{w}(T, X) \cdot X^k \mid T = t\right] - \mathbb{E}[X^k]$ is the *order-k balancing error* corresponding to X^k. We derive this decomposition for binary, discrete, and continuous confounders in Section B.1 This error decomposition suggests a training scheme for learning stabilized weights to minimize all balancing errors upto order K:

$$\hat{w}(T_i, X_i) = \arg\min_{w} \sum_{i=1}^{n} \sum_{k=0}^{K} \|\mathbb{E}[w(T_i, X_i) \cdot X^k] - \mathbb{E}[X^k]\|^2, \quad (10)$$

For IPW-CRM, we add this as a regularization term to the standard propensity score training objective.

3.3. Lower-dimensional projections of high-dimensional APO estimates

How do we use high-dimensional APO estimates in practice? In the case of text treatments, for instance, we may compare APOs of two texts or rank some number of text candidates to select the best one for an outcome of interest. We may also be interested in the effects of some lower dimensional attribute of the text, such as length or sentiment, in order to deduce general strategies that favor an outcome. We denote this attribute as T' and assume it satisfies the following.

Assumption 2. *A known conditional distribution $p_{T'|T}$ relates T' and T: $p_{T'}(t') = \mathbb{E}_T[p_{T'|T}(t'|T)]$. Further, T' is conditionally independent of X and $Y(\cdot)$ given T: $T' \perp (X, Y(\cdot)) \mid T$.*

This assumption holds trivially for deterministic properties of the treatment, for instance, mention of particular words in the text or the length of the text. Under Assumption 2, the lower-dimensional APO $g'(t')$ is given by the expectation of high-dimensional APOs $g(\cdot)$ over samples where $T' = t'$, and can be computed via a Monte Carlo estimate as:

$$g'(t') = \mathbb{E}_{T|T'=t'}[g(T)] \approx \frac{\sum_{i=1}^{n} \mathbb{I}\left(T'_i = t'\right) \hat{g}(T_i)}{\sum_{i=1}^{n} \mathbb{I}\left(T'_i = t'\right)}. \tag{11}$$

See Section B.2 for a complete description of the lower-dimensional projections of treatments and derivation of their APOs. Equation (11) enables the use of a single learned $\hat{g}$ to simultaneously answers causal questions about any lower-dimensional attribute that can be labeled given the high-dimensional treatment, without retraining a separate estimator per attribute.

4. Related Work

High-dimensional and text treatments. Causal inference with complex, multi-valued, or high-dimensional treatments is a well-recognized challenge (Lopez & Gutman, 2017; Wang et al., 2026). Hirano & Imbens (2005) defined the generalized propensity score (GPS) to extend propensity-score-based techniques to continuous treatments. Sharma et al. (2020) studied high-dimensional settings (upto 200 treatments) using bag-of-words representations, with access to multiple observations per treatment value. A growing literature applies these ideas when the treatment is expressed in natural language (Feder et al., 2022), though prior work has largely circumvented the high-dimensionality of text by first extracting lower-dimensional representations (Imai & Nakamura, 2024). Pryzant et al. (2018) learned lexical features as a treatment proxy, Fong & Grimmer (2016) and Pryzant et al. (2020) inferred a latent treatment variable from the text, and Guo et al. (2024) fine-tuned a BERT-based model to predict outcomes from (synthetic) text with underlying binary treatment variables. Zhang et al. (2020) similarly reduced counselor conversations to a handful of attributes (*e.g.* length, speed, sentiment). Wood-Doughty et al. (2018) highlighted the challenges of using text classifiers for such dimensionality reduction. Large language models are also increasingly being used in causal pipelines where text plays roles other than treatment, for structuring patient self-reports (Dhawan et al., 2024) or controlling for it as a confounder (Ma et al., 2025). Our work departs from the above by treating each unique text string as a distinct treatment value. CRM estimators learn APOs in the high-dimensional treatment space and can be projected onto lower-dimensional attributes as needed.

The study of text treatments is motivated by a wide range of applications. In finance, earnings call transcripts have been used to estimate the effect of expressed political risk or sentiment on stock volatility and movement (Zhou & He, 2023; Yang & Shen, 2025), and the effect of COVID-19 exposure language on firm capital structure (Ongsakul et al., 2025; Hassan, 2020). In consumer settings, the causal effect of review sentiment on purchase behavior (Yang & Shen, 2025; Pryzant et al., 2020) and complaint politeness on resolution time (Pryzant et al., 2020) have been studied. In mental health applications, text-treatment APO estimation has potential in evaluating therapist response quality (Bertagnolli, 2020; Ewbank et al., 2020; Althoff et al., 2016) and counseling interventions (Xu et al., 2025). Across all these settings, prior work has relied on manual annotation or classification to reduce treatments to a lower dimensions *a priori*.

Covariate balancing. Correct propensity scores balance functions of covariates across treatments, a property that has inspired several methods to learn balancing weights for binary (Imai & Ratkovic, 2014; Hainmueller, 2012; Wong & Chan, 2018) or continuous (Fong et al., 2018) treatments. Athey et al. (2018) combined first-moment balancing weights with regularized regression adjustment for high-dimensional linear outcome models. Kallus (2020b) developed a generalized optimal matching framework that subsumes covariate balance and establishes a balance-variance trade-off. Other approaches include adversarial balancing (Du et al., 2021; Kallus, 2020a; Ozery-Flato et al., 2018; Kazemi & Ester, 2024), end-to-end stable weight learning (Bahadori et al., 2022), and joint propensity-outcome training with targeted regularization (Shi et al., 2019), all for either binary or continuous treatments. Our contribution connects APO estimation error directly to moment-balancing errors and applies to the high-dimensional discrete treatment setting, which allows the handling of text treatments in the space of possible strings.

5. Empirical Evaluation

Our experiments investigated the following questions: (i) Does balance regularization improve direct APO estimation in a simple linear setting with continuous treatments? (ii) How do the variants of our method perform in synthetic settings with high-dimensional discrete treatments, with varying sample size, model size, confounding strength, and order of moments balanced? (iii) Do our methods scale to more realistic settings to predict potential outcomes of text treatments using large language models? (iv) Can we repurpose a single high-dimensional APO estimator to answer several lower-dimensional causal questions, without training separate estimators for each?

Datasets. We evaluated different methods on three settings of increasing complexity. First, we verified our error decomposition in a simple linear setting with continuous Gaussian

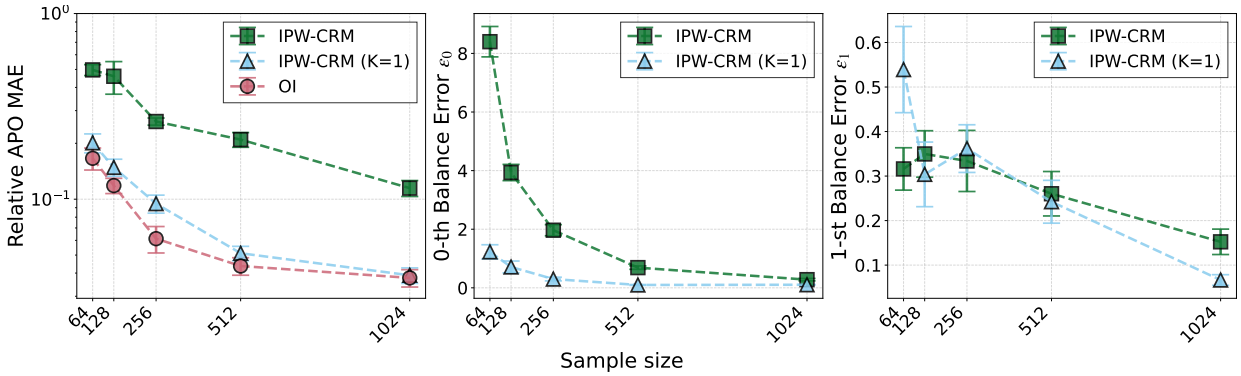

Figure 2. Balancing the order-0 and order-1 moments significantly improved IPW-CRM performance, matching that of OI, for linear models and continuous treatments and covariates.

confounders, treatments, and outcomes, where the true APO is a linear function of confounders and the treatment. Next, we constructed a synthetic setting with discrete confounders, binary outcome, high-dimensional discrete treatments to mirror natural language tokens, and nonlinear relationships between different variables. Finally, we evaluated different methods on a real-world dataset of Amazon Reviews (Hou et al., 2024) in a semi-synthetic setting, where each product review is a different treatment, total number of ratings received by the product is the 8-dimensional confounder, and the purchase of the product by a user reading a given review is the binary outcome of interest. Outcomes were synthetically generated by using GPT-5.1 (Singh et al., 2025) as a conditional outcome model, which we prompted for a realistic probability of purchase, given each possible confounder value and treatment review. All datasets were split into training, validation, and test subsets such that treatments are non-overlapping across the subsets. In the case of Amazon Reviews, each treatment (or review) is observed only once. Complete details for all datasets and experimental parameters are provided in Sections C and D, respectively.

Evaluation. Across all settings we evaluated APO estimation quality via the relative mean absolute error (MAE) and correlation between estimated and true APOs over held-out treatments. For estimators that measure and optimize balance errors, we also report the average of balance errors upto some order K, where K is specific to each setting.

5.1. Does optimizing balance errors reduce APO error in a linear continuous setting?

We used continuous Gaussian data with a linear true APO function to empirically verify the APO error decomposition into the order-0 and order-1 balance errors, as analytically derived in Section C.1. Notice that the continuous setting requires generalization to unseen treatments, as in the case of text treatments. We compared IPW-CRM estimators with and without balance regularization against an oracle outcome imputation estimator (OI) that trains a correctly specified linear network to model the conditional outcomes. Figure 2 shows relative APO MAE, alongside the corresponding order-0 and order-1 balance errors for the IPW-CRM estimators, all of which reduce as sample size increases for all estimators. Without any balance error regularization, IPW-CRM has larger balance errors that persist even as sample size grows, with correspondingly high APO error. In contrast, the balance error regularizer used in IPW-CRM with $K = 1$ substantially reduced balancing errors, which in turn reduced APO error, as expected, and matched the oracle performance of OI. This empirically verifies that minimizing moment-balancing errors directly translates to reduced APO error.

5.2. How does CRM perform with discrete treatments and varying experimental parameters?

As a step towards high-dimensional text treatments, we used a synthetic dataset of high-dimensional discrete treatments and trained small transformer models to compare the performance of different estimators. As shown in Table 1, SW-CRM ($K = 2$) outperforms vanilla OI and other variants. Including higher-order moment-balancing regularization improves its performance. For IPW-CRM, we report results for the best performing $K = 1$. For larger K, IPW-CRM suffered from training instability and hence, poorer performance, a pitfall avoided by SW-CRM.

For the best-performing method, SW-CRM, we conducted further ablation studies to tease apart the effects of different experimental conditions on the correlation between predicted and true APOs, which are visualized in Figure 3. Performance steadily improved with sample size as well as model size, yielding better correlation and lower standard errors. To systematically increase model size, we increased the transformer embedding dimension and its feedforward hidden dimension in proportion. Performance remained stable with increasing confounding strength, suggesting that

Table 1. On both, synthetic and Amazon Reviews, datasets, SW-CRM outperformed other estimators, with higher-order moment balancing improving performance across metrics.

			Synthetic Discrete		Amazon Reviews	
Estimator	K	**Test-time X**	**Rel. MAE**	**APO Correlation**	**Rel. MAE**	**APO Correlation**
OI	–	✓	0.184	0.943	0.167	0.830
OI-CRM	–	✗	0.143	0.969	0.226	0.812
IPW-CRM	1	✗	0.166	0.901	0.871	0.190
SW-CRM	1	✗	0.141	0.935	0.279	0.824
SW-CRM	2	✗	**0.121**	**0.996**	**0.220**	**0.874**

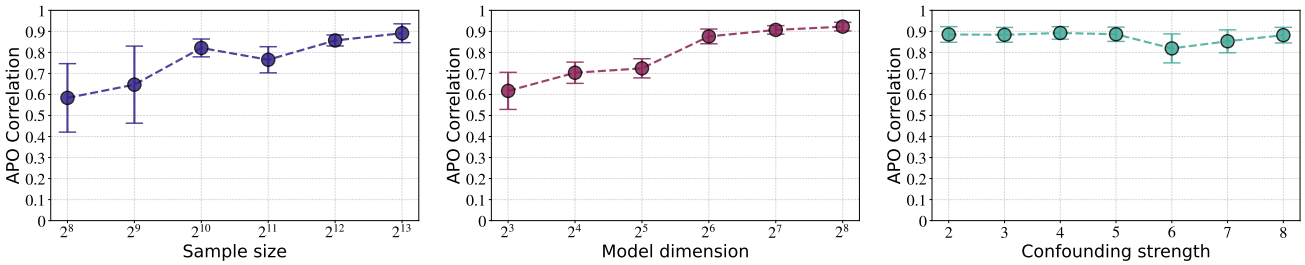

Figure 3. Consistent with standard risk minimization from samples, SW-CRM performance improved with sample size (**Left**) and model size (**Middle**) for the synthetic discrete dataset with transformer models. Given observed confounders, performance remained consistent with the strength of confounding (**Right**), with increase in standard error at higher strengths.

the estimator was able to correct for large amounts of confounding as long as all confounders were known and the assumption of no unmeasured confounding was satisfied. Finally, we studied the impact of the maximum moment order balanced, K, on APO correlation and the total balance error. Figure 4 shows that as moments of higher orders are included in the balance-error regularizer, APO correlation monotonically increased and correspondingly, the total sum of balance errors upto order-4 decreased. This is consistent with our APO error decomposition for non-binary discrete covariates.

5.3. Do our methods scale to text treatments and large language models?

To evaluate our methods with treatments in the very high-dimensional and discrete space of all strings, we used the Amazon Reviews dataset, treating each individual review as a treatment, and finetuned Gemma-3-270M to train different components of our estimators. Note that in this more realistic setting, each treatment is observed in only one datapoint and APO estimation relies on the language model's ability to generalize to unseen treatments. Table 1 shows similar trends in estimators' performance to those of the synthetic dataset. IPW-CRM failed more dramatically in this setting, which is explained by extreme propensity scores in the large treatment space of texts and resulting training instability. SW-CRM avoided this instability and performed competitively with OI and OI-CRM even at $K = 1$, with higher-order moments ($K = 2$) further improving performance. Notably, SW-CRM and OI-CRM achieved high correlations with the true APOs without requiring access to confounder values at test time, which is necessary for inference with standard OI.

5.4. Can we repurpose a single high-dimensional APO estimator to answer several lower-dimensional causal questions efficiently?

Finally, we explored how a single high-dimensional APO estimator may be operationalized and used in practice. A straightforward use-case is to compare estimated APOs of two or more texts to predict which one would lead to the best outcome of interest. It is also possible to ask lower dimensional causal questions about specific attributes of the text and answer them by projecting the high-dimensional APO estimate down from the space of texts to the space of the attribute in question, using Equation (11). For instance, in the case of Amazon Reviews, we maybe interested in the effects of the review rating $(1, \ldots, 5)$, sentiment (positive or negative), or length (greater or less than 100 words) on the outcome of purchase. For each estimator, we computed the projected APOs for these three treatment spaces and compared them against their separately retrained attribute-specific counterparts.

Table 2 shows that the projected estimates from a single high-dimensional APO estimator have competitive or better performance than their retrained counterparts, with SW-CRM

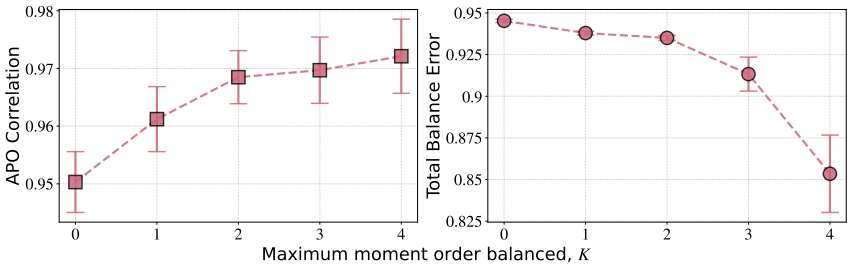

Figure 4. With discrete covariates, balancing higher-order moments directly reduced balance error and hence, APO estimation error, thereby improving correlation with ground-truth.

Table 2. Projecting a single high-dimensional APO estimator for Amazon Reviews onto lower-dimensional treatment attributes (rating, sentiment, length) achieves near-perfect correlation with ground-truth, outcompeting separate estimators that were retrained per attribute.

		Varying true APOs				**Near-constant true APOs**	
		Rating: $\lvert\mathcal{T}\rvert = 5$		**Sentiment:** $\lvert\mathcal{T}\rvert = 2$		**Length:** $\lvert\mathcal{T}\rvert = 2$	
Estimator		**Rel. MAE**	**Correlation**	**Rel. MAE**	**Correlation**	**Rel. MAE**	**Correlation**
OI	Projected	0.118	0.998	0.033	**1.000**	0.091	-1.000
	Retrained	0.359	0.983	0.012	1.000	0.010	1.000
OI-CRM	Projected	0.103	**0.999**	0.038	**1.000**	0.099	-1.000
	Retrained	0.245	0.996	0.021	1.000	0.040	1.000
IPW-CRM (K=1)	Projected	0.778	0.976	0.897	**1.000**	0.903	**1.000**
	Retrained	0.767	0.977	0.510	1.000	0.899	1.000
SW-CRM (K=1)	Projected	0.129	0.991	0.119	**1.000**	0.091	**1.000**
	Retrained	0.091	0.998	0.008	1.000	0.075	1.000
SW-CRM (K=2)	Projected	**0.057**	**0.999**	**0.007**	**1.000**	**0.089**	**1.000**
	Retrained	0.042	0.999	0.006	1.000	0.075	1.000

($K = 2$) outperforming other variants. This performance is consistent across different estimators for the Rating and Sentiment treatments, with no additional training cost for each attribute. The case of Length provides a challenging task with two possible treatments (*i.e.* possible correlations of -1 or 1) with near-constant true APOs. This means that small errors in predictions can reverse the treatment ordering (*i.e.* -1 correlation), which was the case with projected estimates of OI and OI-CRM. Overall, these results show that a single high-dimensional APO estimator can serve as an efficient tool for answering multiple downstream causal queries simultaneously, without retraining separate estimators for each, as long as the assumption of no unmeasured confounding is satisfied and the lower-dimensional attributes satisfy Assumption 2.

6. Limitations and Conclusions

In this work, we recast average potential outcome (APO) estimation with high-dimensional treatments as causal risk minimization (CRM) over the input space of treatments, overcoming challenges of generalization to unseen treatments. We presented an APO error decomposition into moment-balancing errors to learn accurate targets for risk minimization and showed how to efficiently answer multiple causal questions about treatment attributes by projecting a single high-dimensional estimator onto lower-dimensional attributes, with no additional training cost.

Our work has several limitations that may impact its application, including those shared with any observational study: (i) the strong ignorability assumption is necessary, though untestable in practical scenarios; (ii) optimization of higher-order moment-balancing errors translates to improved performance but also incurs additional computational costs that is linear in K; (iii) generalization of CRM estimators relies on inductive biases of the trained model and training data, making it subject to the usual pitfalls of out-of-distribution settings; and (iv) our lower-dimensional projected APOs require access to the relationship between the high-dimensional treatment and its attributes, $p_{T'|T}$, along with the conditional independence assumption that requires T' to depend only on T.

When the required assumptions are satisfied, our results empirically verify the analytic APO error decomposition and demonstrate performance trends with scaled dataset

and model sizes that are consistent with standard risk minimization. The semi-synthetic Amazon Reviews experiments show successful application of CRM estimators to real text treatments and *post hoc* lower-dimensional projections that are competitive with attribute-specific retraining, unlocking the potential for real-world causal estimation in domains with truly high-dimensional treatments.

Acknowledgements

We would like to thank Leonardo Cotta for feedback on a draft of the paper and Rahul Krishnan for helpful discussions. Resources used in preparing this research were provided in part by the Province of Ontario, the Government of Canada through CIFAR, and companies sponsoring the Vector Institute. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), RGPIN-2021-03445.

References

Althoff, T., Clark, K., and Leskovec, J. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics*, 4:463–476, 2016.

Athey, S., Imbens, G. W., and Wager, S. Approximate residual balancing: debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80(4): 597–623, 2018.

Bahadori, T., Tchetgen, E. T., and Heckerman, D. End-to-end balancing for causal continuous treatment-effect estimation. In *International conference on machine learning*, pp. 1313–1326. PMLR, 2022.

Bertagnolli, N. Counsel chat: Bootstrapping high-quality therapy data, 2020. URL https://huggingface.co/datasets/nbertagnolli/counsel-chat.

Capstick, A., Krishnan, R. G., and Barnaghi, P. Autoelicit: Using large language models for expert prior elicitation in predictive modelling, 2025. *URL https://arxiv.org/abs/2411.17284*, 2025.

Choi, K., Cundy, C., Srivastava, S., and Ermon, S. Lmpriors: Pre-trained language models as task-specific priors, 2022. *URL https://arxiv. org/abs/2210.12530*, 2022.

Dai, H., Saccardo, S., Han, M. A., Roh, L., Raja, N., Vangala, S., Modi, H., Pandya, S., Sloyan, M., and Croymans, D. M. Behavioural nudges increase covid-19 vaccinations. *Nature*, 597(7876):404–409, 2021.

Dhawan, N., Cotta, L., Ullrich, K., Krishnan, R. G., and Maddison, C. J. End-to-end causal effect estimation from unstructured natural language data. *Advances in Neural Information Processing Systems*, 37:77165–77199, 2024.

Du, X., Sun, L., Duivesteijn, W., Nikolaev, A., and Pechenizkiy, M. Adversarial balancing-based representation learning for causal effect inference with observational data. *Data Mining and Knowledge Discovery*, 35(4): 1713–1738, 2021.

Ewbank, M. P., Cummins, R., Tablan, V., Bateup, S., Catarino, A., Martin, A. J., and Blackwell, A. D. Quantifying the association between psychotherapy content and clinical outcomes using deep learning. *JAMA psychiatry*, 77(1):35–43, 2020.

Feder, A., Keith, K. A., Manzoor, E., Pryzant, R., Sridhar, D., Wood-Doughty, Z., Eisenstein, J., Grimmer, J., Reichart, R., Roberts, M. E., et al. Causal inference in natural language processing: Estimation, prediction, interpretation and beyond. *Transactions of the Association for Computational Linguistics*, 10:1138–1158, 2022.

Fong, C. and Grimmer, J. Discovery of treatments from text corpora. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1600–1609, 2016.

Fong, C., Hazlett, C., and Imai, K. Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements. *The Annals of Applied Statistics*, 12(1):156–177, 2018.

Graham, R. L., Knuth, D. E., and Patashnik, O. *Concrete mathematics: a foundation for computer science*. Addison-Wesley, 2nd edition, 1994.

Guo, S., Marmarelis, M. G., Morstatter, F., and Lerman, K. Estimating causal effects of text interventions leveraging llms. *arXiv preprint arXiv:2410.21474*, 2024.

Hainmueller, J. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, 20 (1):25–46, 2012.

Hassan, T. A. Firm-level exposure to epidemic diseases: Covid-19, sars, and h1n1 tarek a. hassan, stephan hollander, laurence van lent, and ahmed tahoun. 2020. URL https://www.nber.org/system/files/working_papers/w26971/w26971.pdf.

Hassanpour, N. and Greiner, R. Counterfactual regression with importance sampling weights. In *IJCAI*, pp. 5880–5887. Macao, 2019.

Hirano, K. and Imbens, G. *The Propensity Score with Continuous Treatments*, pp. 73–84. Wiley-Blackwell, July 2005. ISBN 047009043X. doi: 10.1002/0470090456.ch7.

Hirano, K., Imbens, G. W., and Ridder, G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

Holland, P. W. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.

Horvitz, D. G. and Thompson, D. J. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685, 1952.

Hou, Y., Li, J., He, Z., Yan, A., Chen, X., and McAuley, J. Bridging language and items for retrieval and recommendation. *arXiv preprint arXiv:2403.03952*, 2024.

Imai, K. and Nakamura, K. Causal representation learning with generative artificial intelligence: Application to texts as treatments. *arXiv preprint arXiv:2410.00903*, 2024.

Imai, K. and Ratkovic, M. Covariate balancing propensity score. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(1):243–263, 2014.

Jung, Y., Tian, J., and Bareinboim, E. Learning causal effects via weighted empirical risk minimization. *Advances in neural information processing systems*, 33: 12697–12709, 2020.

Kallus, N. Deepmatch: Balancing deep covariate representations for causal inference using adversarial training. In *International Conference on Machine Learning*, pp. 5067–5077. PMLR, 2020a.

Kallus, N. Generalized optimal matching methods for causal inference. *Journal of Machine Learning Research*, 21(62): 1–54, 2020b.

Kazemi, A. and Ester, M. Adversarially balanced representation for continuous treatment effect estimation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 13085–13093, 2024.

Lopez, M. J. and Gutman, R. Estimation of causal effects with multiple treatments: a review and new ideas. *Statistical Science*, pp. 432–454, 2017.

Ma, Y., Frauen, D., Schweisthal, J., and Feuerriegel, S. Llm-driven treatment effect estimation under inference time text confounding. *arXiv preprint arXiv:2507.02843*, 2025.

Milkman, K. L., Patel, M. S., Gandhi, L., Graci, H. N., Gromet, D. M., Ho, H., Kay, J. S., Lee, T. W., Akinola, M., Beshears, J., et al. A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. *Proceedings of the National Academy of Sciences*, 118(20):e2101165118, 2021.

Ongsakul, V., Chatjuthamard, P., Chintrakarn, P., and Jiraporn, P. Leveraging the pandemic: exploring how covid exposure shapes capital structure using a text-based approach. *Review of Behavioral Finance*, 17(2):272–297, 2025.

Ozery-Flato, M., Thodoroff, P., Ninio, M., Rosen-Zvi, M., and El-Hay, T. Adversarial balancing for causal inference. *arXiv preprint arXiv:1810.07406*, 2018.

Pryzant, R., Shen, K., Jurafsky, D., and Wagner, S. Deconfounded lexicon induction for interpretable social science. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pp. 1615–1625, 2018.

Pryzant, R., Card, D., Jurafsky, D., Veitch, V., and Sridhar, D. Causal effects of linguistic properties. *arXiv preprint arXiv:2010.12919*, 2020.

Robins, J. M., Hernan, M. A., and Brumback, B. Marginal structural models and causal inference in epidemiology, 2000.

Rosenbaum, P. R. Model-based direct adjustment. *Journal of the American statistical Association*, 82(398):387–394, 1987.

Rubin, D. B. Assignment to treatment group on the basis of a covariate. *Journal of educational Statistics*, 2(1):1–26, 1977.

Rubin, D. B. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American statistical association*, 75(371):591–593, 1980.

Sharma, A., Gupta, G., Prasad, R., Chatterjee, A., Vig, L., and Shroff, G. Hi-ci: Deep causal inference in high dimensions. In *Proceedings of the 2020 KDD Workshop on Causal Discovery*, pp. 39–61. PMLR, 2020.

Shi, C., Blei, D., and Veitch, V. Adapting neural networks for the estimation of treatment effects. *Advances in neural information processing systems*, 32, 2019.

Singh, A., Fry, A., Perelman, A., Tart, A., Ganesh, A., El-Kishky, A., McLaughlin, A., Low, A., Ostrow, A., Ananthram, A., et al. Openai gpt-5 system card. *arXiv preprint arXiv:2601.03267*, 2025.

Swaminathan, A. and Joachims, T. Counterfactual risk minimization: Learning from logged bandit feedback. In *International conference on machine learning*, pp. 814–823. PMLR, 2015.

Wang, Y., Li, H., Zhu, M., Wu, A., Li, B., Yin, K., Xiong, R., Wu, F., and Kuang, K. Causal inference with complex treatments: A survey. *ACM Computing Surveys*, 58(9): 1–36, 2026.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. L., Gugger, S., Drame, M., Lhoest, Q., and Rush, A. M. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 38–45, Online, October 2020. Association for Computational Linguistics. URL https://www.aclweb.org/anthology/2020.emnlp-demos.6.

Wong, R. K. and Chan, K. C. G. Kernel-based covariate functional balancing for observational studies. *Biometrika*, 105(1):199–213, 2018.

Wood-Doughty, Z., Shpitser, I., and Dredze, M. Challenges of using text classifiers for causal inference. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pp. 4586–4598, 2018.

Xu, J., Wei, T., Hou, B., Orzechowski, P., Yang, S., Jin, R., Paulbeck, R., Wagenaar, J., Demiris, G., and Shen, L. Mentalchat16k: A benchmark dataset for conversational mental health assistance. *arXiv preprint arXiv:2503.13509*, 2025.

Yang, Z. and Shen, B. Estimating textual treatment effect via causal disentangled representation learning. *The Journal of Supercomputing*, 81(2):386, 2025.

Zhang, J., Mullainathan, S., and Danescu-Niculescu-Mizil, C. Quantifying the causal effects of conversational tendencies. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–24, 2020.

Zhou, Y. and He, Y. Causal inference from text: Unveiling interactions between variables. *arXiv preprint arXiv:2311.05286*, 2023.

A. Notation

X	Random variable corresponding to confounders.
T	Random variable corresponding to high-dimensional treatment.
T'	Random variable corresponding to low-dimensional treatment attribute.
Y	Random variable corresponding to outcome observed .
x	Possible instance of X from its support $\mathcal{X}$.
t	Possible instance of T from its support $\mathcal{T}$.
t'	Possible instance of T' from its support.
y	Possible instance of Y from its support $\mathcal{Y} = \{0, 1\}$ or $\mathbb{R}$.
$Y(t)$	Random variable corresponding to potential outcome under treatment t.
X_i	Sampled value of X for individual i.
T_i	Sampled value of T for individual i.
T'_i	Sampled value of T' for individual i.
Y_i	Sampled value of Y for individual i.
$g(t)$	Average potential outcome (APO) given by $\mathbb{E}[Y(t)]$, where the expectation is over some defined population of individuals.
$g'(t')$	Lower-dimensional APO under treatment t'.
$\hat{g}(\cdot)$	Learned APO estimator.
n	Total number of data points.
$\hat{e}(t, x)$	Propensity score estimator, to approximate $p_{T\|X}(t \mid x)$.
$\hat{f}(t, x)$	Conditional outcome estimator given treatment and confounder, to approximate $\mathbb{E}[Y\|T = t, X = x]$.
$\hat{w}(t, x)$	Stabilized weights estimator, to approximate $\frac{p_T(t)}{p_{T\|X}(t\|x)}$.
$\varepsilon_k(t)$	Order-k balancing error.
K	Maximum moment order included in balancing error optimization.

B. Derivations

B.1. APO error decompositions into balancing errors

Recall that under Assumption 1, we can write $f(t, X) = \sum_{k=0}^{K} c_k(t)\, X^k)$ for constants $c_k(t)$. Then, using Equation (8), we have:

$$\hat{g}(t) - g(t) = \sum_k c_k(t) \cdot \varepsilon_k(t), \text{ where } \varepsilon_k(t) = \mathbb{E}\left[\hat{w}(T, X) \cdot X^k \mid T = t\right] - \mathbb{E}[X^k], \tag{12}$$

is the *order-k balancing error* corresponding to X^k. We now instantiate this decomposition for confounders in different settings: binary, discrete, and continuous.

Binary X (warm-up). If $X \in \{0, 1\}$, we have $f(t, X) = f(t, 0) \cdot 1 + (f(t, 1) - f(t, 0)) \cdot X$. Substituting into Equation (12), the APO error depends on two balance error terms:

$$\hat{g}(t) - g(t) = f(t, 0) \cdot \varepsilon_0(t) + \left[f(t, 1) - f(t, 0)\right] \cdot \varepsilon_1(t), \tag{13}$$

where $\varepsilon_0(t) = \mathbb{E}[\hat{w} \mid T = t] - 1$ is a normalization error and $\varepsilon_1(t) = \mathbb{E}[\hat{w}(t, X) \cdot X \mid T = t] - \mathbb{E}[X]$ is the common first moment balance error. In this case, the APO error of the risk minimizer vanishes whenever the weights are properly normalized and balance the first moment of X.

Discrete finite X. For $X \in \{0, 1, \ldots, K\}$, the Newton forward difference formula (Graham et al., 1994), $f(t, \cdot)$ admits the finite expansion,

$$f(t, X) = \sum_{k=0}^{K} \frac{\Delta^k f(t, 0)}{k!} \prod_{j=0}^{k-1} (X - j), \tag{14}$$

where $\Delta^k f(t, 0) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} f(t, j)$ is the k-th forward difference at zero and the product $\prod_{j=0}^{k-1}(X - j)$ is also known as the falling factorial. Notice that we may collect terms to rewrite this expansion in terms of powers of X: $f(t, X) = \sum_{k=0}^{K} c_k(t) X^k$ for some $c_k(t)$. Hence,

$$\hat{g}(t) - g(t) = \sum_{k=0}^{K} c_k(t) \cdot (\mathbb{E}[\hat{w}(t, X) \cdot X^k \mid T = t] - \mathbb{E}[X^k]). \tag{15}$$

In other words, balancing the moments $\mathbb{E}[X^k]$ for $k = 0, \ldots, K$ drives the APO error to zero.

Continuous X. Finally, when $X \in \mathbb{R}$ and f is a continuous polynomial of order at most K, the Taylor expansion of the true conditional outcome function $f(t, x) = \mathbb{E}[Y \mid T = t, X]$ around $X = 0$ yields for each t:

$$f(t, X) = \sum_{k=0}^{K} \frac{f^{(k)}(t, 0)}{k!} X^k \implies g(t) = \mathbb{E}[f(t, X)] = \sum_{k=0}^{K} \frac{f^{(k)}(t, 0)}{k!} \mathbb{E}[X^k], \tag{16}$$

where $f^{(k)}(t, 0) = \frac{\partial^k f(t,x)}{\partial x^k}\big|_{x=0}$. Similarly, the estimated APO is:

$$\hat{g}(t) = \mathbb{E}[\hat{w}(T, X) \cdot f(t, X) \mid T = t] = \sum_{k=0}^{\infty} \frac{f^{(k)}(t, 0)}{k!} \mathbb{E}\left[\hat{w}(T, X) \cdot X^k \mid T = t\right]. \tag{17}$$

Let the order-k balancing error be $\varepsilon_k(t) = \mathbb{E}[\hat{w}(T, X) \cdot X^k \mid T = t] - \mathbb{E}[X^k]$. Then,

$$\hat{g}(t) - g(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(t, 0)}{k!} \cdot \varepsilon_k(t). \tag{18}$$

Across all cases, the APO estimation error of the risk minimizer can be written in terms of moment balancing errors of the confounders.

B.2. Deriving lower-dimensional projections of high-dimensional APOs

Let T' denote the lower-dimensional attribute of the high-dimensional treatment T, where the relationship between the two variables is specified through a known conditional distribution $p_{T'|T}$, so that $p_{T'}(t') = \mathbb{E}_T[p_{T'|T}(t'|T)]$. We assume T' is independent of X and $Y(\cdot)$ given T.

To relate the APOs $g(\cdot)$ over the support of T to the APOs $g'(\cdot)$ over the support of T', consider the ideal completely randomized experiment (CRE) where the assignment distribution over T induces an assignment distribution over T'. The randomized distribution over all variables is given by

$$p_{\text{cre}}(y, x, t, t') = p_{Y(t),X}(y, x)p_T(t)p_{T'|T}(t'|t).$$

Under this distribution, we can estimate joint-treatment APOs under $T = t$ and $T' = t'$ as:

$$\begin{aligned}
g(t, t') &= \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i \mathbb{I}\Big(T_i = t, T_i' = t'\Big)}{p_{T,T'}(T_i, T_i')} \\
&\approx \sum_{x,s,s'} \int \frac{y\mathbb{I}\Big(s = t, s' = t'\Big)}{p_{T,T'}(s, s')} p_{\text{cre}}(y, x, s, s')\, dy \\
&= \mathbb{E}_{X,Y|T=t,T'=t'}[Y|T = t, T' = t'] \\
&= \mathbb{E}_{X,Y|T=t}[Y|T = t] \\
&= g(t),
\end{aligned}$$

where the second-to-last equality follows from the conditional independence of T' and $(X, Y(\cdot))$ given T. Hence, under the randomized distribution over T', we have APOs given by:

$$\begin{aligned}
g'(t') &= \mathbb{E}_{X,Y|T'=t'}[Y|T' = t'] \\
&= \mathbb{E}_{T|T'=t'}[\mathbb{E}_{X,Y|T'=t',T}[Y|T' = t', T] \\
&= \mathbb{E}_{T|T'=t'}[g(T, t')] \\
&= \mathbb{E}_{T|T'=t'}[g(T)].
\end{aligned}$$

Our experiments estimate g with different APO estimators $\hat{g}$ and the expectation above with the Monte Carlo estimate:

$$\hat{g}'(t') = \frac{\sum_{i=1}^{n} \mathbb{I}\Big(T_i' = t'\Big)\, \hat{g}(T_i)}{\sum_{i=1}^{n} \mathbb{I}\Big(T_i' = t'\Big)}.$$

C. Dataset Details

C.1. Linear Continuous Setting

We generated data according to:

$$\begin{aligned} X_i &\sim \mathcal{N}(0,1), \\ T_i \mid X &\sim \mathcal{N}(X,1), \\ Y_i &= 1 + 2T + 3X + \epsilon, \end{aligned}$$

where $\epsilon \sim \mathcal{N}(0,1)$ is independent noise. The true APOs are given by $g(t) = 1 + 2t + 3\,\mathbb{E}[X]$.

The APO error decomposition into the order-0 and order-1 moment-balancing error follows from the continuous case analyzed in Section 3.2, or can be directly derived for this data generating process.

$$\begin{aligned} \hat{g}_{\text{SW-CRM}}(t) - g(t) &= \mathbb{E}[\hat{w}(T,X)Y \mid T=t] - g(t) \\ &= \mathbb{E}[\hat{w}(T,X) \mid T{=}t] + 2t\mathbb{E}[\hat{w}(T,X) \mid T=t] \\ &\quad + 3\mathbb{E}[\hat{w}(T,X)X \mid T=t] + \mathbb{E}[\hat{w}(T,X)\epsilon \mid T=t] \\ &\quad - (1 + 2t + 3\,\mathbb{E}[X]) \\ &= (1+2t)(\mathbb{E}[\hat{w}(T,X) \mid T{=}t] - 1) + 3(\mathbb{E}[\hat{w}(T,X)X \mid T=t] - \mathbb{E}[X]) \\ &\quad + \mathbb{E}[\hat{w}(T,X) \mid T=t]\,\mathbb{E}[\epsilon] \\ &= (1+2t)\varepsilon_0(t) + 3\varepsilon_1(t). \end{aligned}$$

C.2. Synthetic Discrete Dataset

We generated discrete sequence data with length-4 confounders and length-3 treatments according to:

$$\begin{aligned} X_i^{(k)} &\sim \text{Uniform}\big(0,\ldots,V_k^X{-}1\big), \quad k=0,1,2,3, \\ T_i^{(j)} \mid X_i &\sim \text{Categorical}\big(p_j(\cdot \mid X_i)\big), \quad j=0,1,2, \\ Y_i \mid T_i, X_i &\sim \text{Bernoulli}\big(\sigma(\mu(T_i,X_i))\big), \end{aligned}$$

where "vocabulary" sizes were $V^X = (5,4,2,3)$, such that $|\mathcal{X}| = 120$, and $V^T = (4,2,2)$, such that $|\mathcal{T}| = 16$. For the model size ablation study, we used $V^T = (16,2,2)$, such that $|\mathcal{T}| = 64$, to demonstrate the benefits of larger models. The true propensity factorizes across the treatment sequence, $p_{T|X}(t \mid x) = \prod_{j=0}^{2} p_j(t^{(j)} \mid x)$, where

$$p_j(t^{(j)} \mid x) = \frac{\exp(\ell_j(x, t^{(j)}))}{\sum_{t^{(j)}=0}^{V_j^T-1} \exp(\ell_j(x, t^{(j)}))},$$

and the logits are

$$\begin{aligned} \ell_0(x, t^{(0)}) &= t^{(0)}\big(x^{(0)} + (x^{(0)})^2 + 2(x^{(0)})^3\big), \\ \ell_1(x, t^{(1)}) &= -t^{(1)}\big(x^{(3)} + (x^{(3)})^2 + 2(x^{(3)})^3\big), \\ \ell_2(x, t^{(2)}) &= t^{(2)}\big(x^{(1)} + (x^{(1)})^2 + 2(x^{(1)})^3\big). \end{aligned}$$

The true conditional outcome model is

$$\begin{aligned} \mu(t,x) = {} & 0.3t^{(0)} + 0.2t^{(1)} + 0.15t^{(2)} \\ & + 0.4x^{(0)} + 0.10x^{(1)} + 0.25x^{(2)} + 0.15x^{(3)} \\ & + t^{(0)}x^{(0)} + t^{(0)}(x^{(0)})^2 + t^{(0)}(x^{(0)})^3. \end{aligned}$$

The true APOs are computed by exact enumeration:

$$g(t) = \mathbb{E}[Y(t)] = \frac{1}{120} \sum_{x \in \mathcal{X}} \sigma(\mu(t,x)).$$

C.3. Amazon Reviews Text Dataset

We constructed a semi-synthetic dataset from the Amazon Reviews 2023 corpus (Hou et al., 2024), Electronics category, with $n = 10,000$ reviews. The treatment T is the full text of a product review, along with the rating given. The confounder $X \in 0, \ldots, 7$ is the product's total rating count, discretized into 8 bins with thresholds $(10, 50, 200, 1000, 5000, 20000, 100000)$. The outcome was the binary variable corresponding to whether or not a customer purchases the product after reading a given review. Since counterfactual outcomes are unobservable, we used GPT-5.1 to generate the true conditional purchase probabilities for the purpose of evaluation. For each review and each rating count bin, we queried GPT-5.1 with the system prompt:

```
You are a customer shopping for an electronic product online on Amazon.
You get to see the number of ratings the product has and one complete
user review, with their rating.  Based on both these things, what is
the probability (0.0 to 1.0) that you will purchase the product?  Reason
about how the number of ratings and the particular review affect your
probability of purchase:

Larger number of ratings indicate a popular product, making you more
likely to purchase it.
The more positive the review sentiment is, the more likely you are to
purchase it.
A detailed and informative review makes your purchase probability more
aligned with the review sentiment.
A large number of ratings also reinforces the review sentiment and
accordingly your purchase probability.
Respond with only a single float between 0.0 and 1.0, nothing else.
```

and a user prompt of the form:

```
Number of Ratings:  <bin label>
Rating:  <rating>/5
Review:  <review text>
```

yielding $\mu(t,k) = P(Y{=}1 \mid T{=}t, X{=}k)$ for each (t,k) pair. The observed outcome was sampled as $Y_i \sim \text{Bernoulli}(\mu(T_i, X_i))$. The true APO was computed as:

$$g(t) = \mathbb{E}[Y(t)] = \sum_{k=0}^{7} P(X{=}k)\, \mu(t,k),$$

where $P(X{=}k)$ is the empirical bin frequency.

For the lower-dimensional projections of text APOs, we annotated each review with the following three attributes:

1. Rating: the rating out of 5, e.g. `"4/5"`,
2. Sentiment: `"very positive"` if rating > 4, else `"negative or neutral"`,
3. Length: `"more than 100 words"` or `"less than 100 words"`.

D. Experimental Details

All hyperparameters were chosen on a validation split for each dataset and model.

D.1. Linear Continuous Setting

We used linear networks for the conditional outcome model and the APO model that maps treatments to APO targets, each trained with the Adam optimizer at learning rate 10^{-2} for 2000 steps. The propensity model was a 3-layer multilayer perceptron with hidden dimension 64 and ReLU activations, mapping $X \in \mathbb{R}$ to two scalar outputs $\mu_\theta(X)$ and $\log \sigma_\theta(X)$, which parameterized the Gaussian propensities $T \mid X \sim \mathcal{N}(\mu_\theta(X), \sigma_\theta(X)^2)$. It was trained via maximum likelihood, with and without balance-error regularization, using Adam with learning rate 10^{-3} for 2000 steps.

D.2. Synthetic Discrete Dataset

All trained models (propensity, stabilized weights, conditional outcome, and APO) were 2-layer transformers with embedding dimension 8, 2 attention heads, feedforward dimension 32, trained with AdamW and learning rate 10^{-4} on $n = 10,000$ samples. Propensity and stabilized models were trained for 2000 epochs each and APO and outcome models for 3000 epochs with the cross-entropy loss derived in Section 3.1 for binary outcomes.

D.3. Amazon Reviews Text Dataset

For all models, we finetuned a pretrained Gemma-3-270M model with AdamW on their corresponding objectives, using the Transformers library (Wolf et al., 2020) on 2 NVIDIA L40 GPUs. Hyperparameters vary per estimator and are summarized in Table 3.

Table 3. Hyperparameters for Amazon Reviews experiments.

Estimator	Model	LR	Epochs	Batch size
OI	Outcome	10^{-5}	5	16
OI-CRM	Outcome	10^{-5}	10	16
	APO	10^{-5}	10	16
IPW-CRM	Propensity	10^{-6}	10	4
	APO	10^{-6}	10	16
SW-CRM	SW	10^{-5}	10	16
	APO	10^{-5}	5	16