OUTSTRUMENTAL VARIABLE REGRESSION

Anonymous authors
Paper under double-blind review

ABSTRACT

Estimating the causal effect of a treatment X on an outcome Y from observational data in the presence of unobserved confounding requires additional structure or assumptions. Existing identification strategies such as instrumental variable regression exploit exogenous variation in only the treatment . We show that the causal effect can also be identified if the outcome mechanism varies across a finite set of environments, even if the treatment distribution is unchanged. Concretely, we introduce *Outstrumental Variable Regression* (OVR), which proceeds in three steps: (i) compute the conditional expectation of Y given X in each environment; (ii) form the vector of differences in these conditional expectations between all environments and a reference environment (we call this vector *contrast*); and (iii) regress Y on X while adjusting for the contrast and environment. We prove that this yields the true causal effect if the following two conditions hold: *inclusion restriction*, i.e., the contrast determines the conditional confounding bias; and *positivity*, i.e., the treatment varies sufficiently strongly after adjustment. We prove that a finite-sample version of OVR is consistent under regularity conditions and demonstrate the procedure's effectiveness on synthetic and real-world data.

1 INTRODUCTION

Estimating the causal effect of a treatment X on an outcome Y is a central goal of empirical science. Randomized controlled trials (Fisher, 1935) provide a direct solution, but are often infeasible, unethical, or too costly. When only observational data is available, unobserved confounding usually renders causal effects unidentifiable without additional assumptions.

One of the most widely used approaches to this problem is instrumental variable (IV) regression, which has played a central role in estimating causal effects from natural experiments and policy changes (Angrist et al., 1996; Angrist & Pischke, 2009). In graphical terms, an instrument E supplies a 'treatment channel': it generates variation in X (via $E \to X$ in Fig. 1), while the exclusion restriction rules out alternative pathways from E to Y, captured in Fig. 1 as $E \not\to H$ and $E \not\to Y$ (Pearl, 2009; Peters et al., 2017).

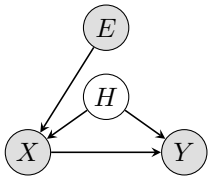


Figure 1: IV setting.

We propose a complementary identification approach that exploits a different source of information: variation in the 'outcome channel', that is, $E \to Y$ (see Fig. 2). Assume E is categorical and corresponds to different environments in which we observe the data (for example, different sites, time periods, experimental conditions, or policy regimes). The intuition behind our method is as follows: because the causal effect is invariant, changes in the X–Y relationship across environments (the contrasts) can reveal what the confounding contribution looks like (i.e., the 'shape' of the confounding). As we will see, we can therefore remove the confounding by adjusting for the contrast and the environment. More precisely, our method proceeds in three steps (see Fig. 3 for an overview). First, we estimate $\mu_e(x) = \mathbb{E}[Y|X = x, E = e]$. Second, we compute a vector of differences between these conditional expectations relative to an arbitrary reference environment e_0, $C(x) := (\mu_e(x) - \mu_{e_0}(x))_{e \neq e_0}$. Under the assumption that the causal effect is invariant across environments, its contribution cancels in these differences. What remains therefore carries information about the confounding that biases the observational relationship between X and Y. Third, we use $C(X)$, together with the environment E, as an adjustment variable. We call this procedure *Outstrumental Variable Regression* (OVR).

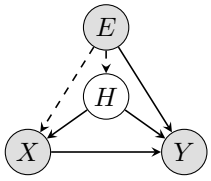


Figure 2: OVR setting.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

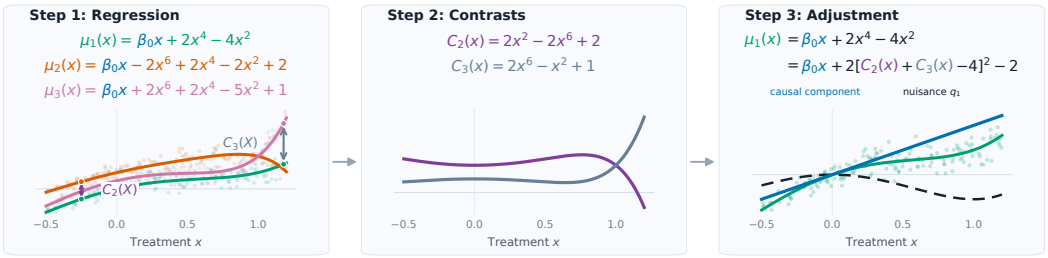


Figure 3: Outstrumental Variable Regression (OVR). We illustrate the main steps of OVR using three environments. **Step 1: Regression.** Estimate the conditional expectations μ_1, μ_2, μ_3 of Y given X in each environment. **Step 2: Contrast.** Subtract μ_1 from μ_2 and μ_3 to obtain the contrasts C_2 and C_3; the shared causal contribution $\beta_0 x$ cancels. **Step 3: Adjustment.** Regress Y on X, adjusting for the contrast vector $C(X) := (C_2(X), C_3(X))$ and environment E. This recovers β_0 if $C(X)$ determines the nuisance $q_1(x) := \mu_1(x) - \beta_0 x$ (dashed black) and enough variation in X remains after partialling out $C(X)$ and E.

Similar to IV, our identifying assumptions, Assumption 1 and Assumption 2, are not primarily graphical restrictions on which causal paths may or may not be present. Instead, they constrain how the outcome mechanism varies across environments. This allows us to leave much of the remaining data-generating process unrestricted, without specifying structural equations for all variables. Our identifying assumptions are thus not implied by the graph in Fig. 2 alone; in § 4, we give three concrete structural settings in which they hold.

OVR is complementary to multi-environment methods that exploit environment-induced variation in the treatment (Rothenhäusler et al., 2019; 2021; Schur et al., 2026b; Gnecco et al., 2026). Thus, OVR can identify the causal coefficient even when the treatment distribution is unchanged across environments. Conversely, IV can identify effects in settings where the outcome contrasts are uninformative. Neither set of identifying conditions contains the other.

Contributions. We establish three results. First, under the inclusion restriction (Assumption 1) and positivity (Assumption 2) the causal effect is identified by OVR (Theorem 1). Second, we highlight three structural settings which imply both of these assumptions, and provide a sensitivity bound that quantifies the bias when the inclusion restriction does not hold (§ 4 and § 3.1). Third, adding Gaussian noise to smooth the empirical version of the contrasts estimator yields a consistent estimator (Theorem 2), which we evaluate on six synthetic settings and a real-data benchmark (§ 6).

2 Causal Reasoning via Outstrumental Variable Regression

We first formalize the problem setting (§ 2.1) and build intuition through an illustrative example (§ 2.2). We then describe the proposed approach for the general case (§ 2.3).

2.1 Problem Setting

Setting. We observe realizations of the random variables $(X, Y, E) \sim P$, where $X \in \mathbb{R}^d$ is a (possibly multivariate) treatment, $Y \in \mathbb{R}$ is an outcome, and $E \in \mathcal{E}$ is an environment label. Throughout, we assume that $\mathcal{E}$ is a finite set. We assume that the average causal effect of X on Y is linear and invariant across environments; more precisely, there exists some $\beta_0 \in \mathbb{R}^d$ and environment-specific intercepts $(\alpha_e)_{e \in \mathcal{E}}$ such that for every $e \in \mathcal{E}$ and P_X-almost all x,

$$\mathbb{E}[Y \mid \mathrm{do}(X = x), E = e] = \beta_0^\top x + \alpha_e. \tag{1}$$

Furthermore, we assume environment overlap: for P_X-almost every x and every $e \in \mathcal{E}$,

$$P(E = e \mid X = x) > 0. \tag{2}$$

We also assume $\mathbb{E}\|X\|_2^2 + \mathbb{E}[Y^2] < \infty$ to ensure that second moments are finite.

Goal. Our goal is to identify β_0 from the (observational) joint distribution P and estimate it from an i.i.d. sample of P. We do not assume access to interventional data. With unobserved con-

founders H, the expression in Eq. (1) generally differs from the observational conditional expectation $\mathbb{E}[Y|X=x,E=e]$. Simply regressing Y on X (after stratifying or pooling environments) thus generally yields an inconsistent estimator of β_0.

Graphical perspective. The considered problem setting includes, for example, data-generating processes captured by a structural causal model (SCM; Pearl, 2009; Bongers et al., 2021) whose induced graph is shown in Fig. 2 (where we allow for the dashed arrows) and with structural equation for Y given by

$$Y := \beta_0^\top X + \alpha_E + \gamma(H, E, X, U_Y), \tag{3}$$

where U_Y is unobserved noise and, for every fixed x, $\mathbb{E}[\gamma(H,E,x,U_Y)|E] = 0$ almost surely. Thus, our setting requires neither the treatment distribution nor the confounder distribution to be invariant across environments.

2.2 An illustrative example

We first illustrate the main idea behind our identification strategy through a simple example with two environments $\mathcal{E} = \{1,2\}$ satisfying $P(E=1), P(E=2) > 0$. Assume X, Y, and H are real-valued and centered and consider the structural model

$$Y := \beta_0 X + \lambda_E H + U_Y, \tag{4}$$

where U_Y satisfies $\mathbb{E}[U_Y|X,E]=0$. We assume $\lambda_1 \neq \lambda_2$, so the environment influences the effect of H on Y (via λ_E). But we assume the environment does not change the joint distribution of (H,X): that is, we have $E \perp\!\!\!\perp (H,X)$, corresponding to the solid-arrow subgraph of Fig. 2. This is a special case of (3) with $\alpha_e = 0$ and $\gamma(H,e,x,U_Y) = \lambda_e H + U_Y$, and hence satisfies (1) without environment-specific intercepts. In population form, our approach consists of three steps.

Step 1: Regression. Compute the conditional expectation of Y given $X=x$ in environment e,

$$\mu_e(x) := \mathbb{E}[Y \mid X=x, E=e] \overset{(4)}{=} \beta_0 x + \lambda_e \mathbb{E}[H \mid X=x], \tag{5}$$

where we have used $\mathbb{E}[U_Y|X,E]=0$ and $H \perp\!\!\!\perp E \mid X$.

Step 2: Contrast. Form the contrast in the means from (5) across environments,

$$C(x) := \mu_2(x) - \mu_1(x) = (\lambda_2 - \lambda_1)\mathbb{E}[H \mid X=x]. \tag{6}$$

The invariant causal effect $\beta_0 x$ cancels from the contrast $C(x)$, while the conditional confounding signal $\mathbb{E}[H|X=x]$ that biases the observational conditional expectation is retained.

Step 3: Adjustment. Use $C(X)$ and E for adjustment. To do so, define the residualized variables

$$\widetilde{X} := X - \mathbb{E}[X \mid C(X), E] \qquad \text{and} \qquad \widetilde{Y} := Y - \mathbb{E}[Y \mid C(X), E]. \tag{7}$$

Taking conditional expectations of (4) given $(C(X), E)$ yields

$$\mathbb{E}[Y \mid C(X), E] = \beta_0 \mathbb{E}[X \mid C(X), E] + \lambda_E \mathbb{E}[H \mid X], \tag{8}$$

where we used that $E \perp\!\!\!\perp (H,X)$ and $\mathbb{E}[H|C(X)] = \mathbb{E}[H|X]$ by (6), as well as $\mathbb{E}[U_Y|C(X),E] = 0$ due to the tower property and $\mathbb{E}[U_Y|X,E] = 0$. Subtracting (8) from (4) then yields

$$\widetilde{Y} = \beta_0 \widetilde{X} + \lambda_E \big(H - \mathbb{E}[H \mid X]\big) + U_Y.$$

Partialling out $C(X)$ and E thus removes exactly the part of the confounding that is predictable from X. The last two terms have a conditional expectation of zero given (X,E). Since $\widetilde{X}$ is a function of (X,E), their product with $\widetilde{X}$ also has an expectation of zero by the tower property. Multiplying by $\widetilde{X}$ and taking expectations yields $\mathbb{E}[\widetilde{X}\widetilde{Y}] = \beta_0 \mathbb{E}[\widetilde{X}^2]$. Consequently, whenever $\mathbb{E}[\widetilde{X}^2] > 0$, the causal parameter β_0 is identified by $\mathbb{E}[\widetilde{X}\widetilde{Y}]/\mathbb{E}[\widetilde{X}^2]$. The contrast $C(X)$ can therefore be used to identify β_0 as long as sufficient variation in X remains after partialling out $C(X)$ and E.

2.3 Outstrumental Variable Regression (OVR)

We now define the general population-level procedure for the setting from § 2.1. For each environment $e \in \mathcal{E}$, define $\mu_e(x) := \mathbb{E}[Y|X = x, E = e]$ as in (5) and the corresponding *nuisance* $q_e(x) := \mu_e(x) - \beta_0^\top x$. Thus, $\mu_e(x) = \beta_0^\top x + q_e(x)$, where q_e contains the environment-specific intercept α_e from (1) and the difference between observational and interventional conditional expectations. (This decomposition does not require specifying an underlying SCM for how (E, X) and potential hidden confounders are generated.) Fix an arbitrary reference environment $e_0 \in \mathcal{E}$ and define the *contrast* C by

$$C(x) := \big(\mu_e(x) - \mu_{e_0}(x)\big)_{e \neq e_0} = \big(q_e(x) - q_{e_0}(x)\big)_{e \neq e_0}. \tag{9}$$

The invariant causal term $\beta_0^\top x$ cancels from each component of $C(x)$. *Outstrumental Variable Regression* (OVR) adjusts for $C(X)$ and E by residualizing with respect to $(C(X), E)$. With $\widetilde{X}$ and $\widetilde{Y}$ defined as the residualized variables from Eq. (7), let

$$\Sigma_C := \mathbb{E}\left[\widetilde{X}\widetilde{X}^\top\right] \qquad \text{and} \qquad \beta_{\text{OVR}} := \Sigma_C^{-1}\mathbb{E}\left[\widetilde{X}\widetilde{Y}\right], \tag{10}$$

whenever Σ_C is nonsingular (Assumption 2). For finite samples, to reduce sensitivity to contrast estimation errors, we introduce Smoothed Outstrumental Variable Regression (Algorithm 1) and prove its consistency in § 3.2.

3 Identifiability and a Consistent Algorithm

We first give sufficient conditions for identifiability under which β_{OVR} equals β_0 in § 3.1 and then establish consistency of the proposed Smoothed OVR estimator in § 3.2.

3.1 Identifiability

Our identification strategy for the invariant causal effect from (1) rests on two key assumptions. First, we require the baseline nuisance $q_{e_0}(X)$ to depend on X only through the contrast $C(X)$. Since the contrasts give the differences $q_e(X) - q_{e_0}(X)$, the full nuisance $q_E(X)$ is then determined by $(C(X), E)$. Residualizing X and Y with respect to $(C(X), E)$ therefore eliminates the nuisance term from the outcome equation. This is what happens in § 2.2: there, $C(X) = (\lambda_2 - \lambda_1)\mathbb{E}[H|X]$, so $\lambda_2 \neq \lambda_1$ makes the contrast determine $\mathbb{E}[H|X]$ and hence the baseline nuisance $q_1(X) = \lambda_1\mathbb{E}[H|X]$.

Assumption 1 (Inclusion restriction (IR)). *There exists a measurable function φ such that the baseline nuisance satisfies $q_{e_0}(X) = \varphi(C(X))$ almost surely.*

If the nuisance q_e does not vary across environments, IR fails unless the nuisance is almost surely constant. Validity of IR does not depend on the choice of reference environment because, for all e, $q_e(X) = q_{e_0}(X) + C_e(X)$. Furthermore, the contrasts corresponding to two different choices of e_0 contain the same information, so the population algorithm does not change, either. Fig. 3 illustrates IR with reference environment $e_0 = 1$: the two contrasts jointly determine the baseline nuisance q_1 through a nonlinear function (Fig. 3, right). Neither contrast alone determines q_1, and no affine function of both contrasts can represent it.

Second, we require that enough variation in X remains after conditioning on $(C(X), E)$.

Assumption 2 (Positivity). $\Sigma_C = \mathbb{E}[\widetilde{X}\widetilde{X}^\top]$ *is positive definite.*

Since, for all $v \in \mathbb{R}^d$, we have $v^\top \Sigma_C v = \mathbb{E}[\text{Var}(v^\top X|C(X), E)]$, positivity means that every nonzero linear combination of X has positive residual variance after conditioning on $(C(X), E)$. Thus, the constrast is required to determine the nuisance while leaving residual variation in every linear treatment direction. If C is injective on the support of X, then $C(X)$ determines X: IR holds automatically, but $\widetilde{X} = 0$ almost surely and $\Sigma_C = 0$, thus violating positivity. Positivity can be assessed from the observed data once C has been estimated, since it concerns the residual covariance of X. IR is different in this regard: it requires the contrasts to capture the unobserved confounding term and therefore cannot be verified from observational data alone.

Theorem 1 (Identification). *Under the setting of § 2.1, Assumptions 1 and 2 imply $\beta_{\text{OVR}} = \beta_0$.*

Algorithm 1 Smoothed Outstrumental Variable Regression

Input: Data $\{(X_i, Y_i, E_i)\}_{i=1}^n$, folds K, smoothing scale τ, number of perturbations M.

1. **Estimate contrasts.** Split data into $\mathcal{I}_C$ and $\mathcal{I}_\beta$, then split $\mathcal{I}_C$ into $\mathcal{I}_A$ and $\mathcal{I}_B$. Fit contrasts $\widehat{C}^A$ and $\widehat{C}^B$ on these respective samples and set $\overline{C} := (\widehat{C}^A + \widehat{C}^B)/2$ and $D := \widehat{C}^A - \widehat{C}^B$.
2. **Fit residualization regressions.** Partition $\mathcal{I}_\beta$ into folds $I_1, \ldots, I_K$. For each k, use $I_{-k} := \mathcal{I}_\beta \setminus I_k$ to estimate $\widehat{\Omega}_k := \widehat{\mathrm{Cov}}_{I_{-k}}(D(X))/4$ and fit regressions $\widehat{f}_{X,k}$ and $\widehat{f}_{Y,k}$ of X and Y on $(\overline{C}(X)+\xi, E)$, where we draw $\xi \sim N(0, \tau^2\widehat{\Omega}_k)$ independently for each training observation.
3. **Partial out on held-out folds.** For each k, draw $\xi'_{im} \sim N(0, \tau^2\widehat{\Omega}_k)$ independently over $i \in I_k$ and $m = 1, \ldots, M/2$. Using the same perturbations for X and Y, set

$$\widehat{Z}_i := Z_i - \frac{1}{M}\sum_{m=1}^{M/2}\sum_{s\in\{-1,1\}} \widehat{f}_{Z,k}\big(\overline{C}(X_i) + s\xi'_{im}, E_i\big) \qquad \text{for} \qquad Z \in \{X, Y\}.$$

Return: $\widehat{\beta}_{\mathrm{SOVR}} := \left(\sum_{i\in\mathcal{I}_\beta} \widehat{X}_i\widehat{X}_i^\top\right)^{-1} \sum_{i\in\mathcal{I}_\beta} \widehat{X}_i\widehat{Y}_i$.

In particular, β_0 is identifiable from the (observational) joint distribution P of (X, Y, E). Proofs of the main-text results are collected in § D.

Relaxing IR. IR, alongside positivity, is sufficient but not necessary for $\beta_{\mathrm{OVR}} = \beta_0$. Define the nuisance remaining after adjustment by $r_E(X) := q_E(X) - \mathbb{E}[q_E(X)|C(X), E]$. Under positivity, $\beta_{\mathrm{OVR}} = \beta_0$ if and only if $\mathbb{E}[\widetilde{X}r_E(X)] = 0$. IR implies the stronger condition $r_E(X) = 0$ almost surely. More precisely, in § D.2 we prove that $\beta_{\mathrm{OVR}} - \beta_0 = \Sigma_C^{-1}\mathbb{E}[\widetilde{X}r_E(X)]$.

Robustness analysis. With a similar reasoning, we can bound the bias of OVR when IR fails to hold. Suppose $q_{e_0}(X) \in L^2$ and define the strength of IR violation

$$\delta_{\mathrm{IR}}^2 := \inf_\varphi \mathbb{E}\Big[\big(q_{e_0}(X) - \varphi(C(X))\big)^2\Big] = \mathbb{E}[\mathrm{Var}(q_{e_0}(X) \mid C(X))]\,,$$

where the infimum is over measurable functions φ. Thus, δ_{IR} measures the variation in the baseline nuisance that is not explained by the contrasts, and IR holds if and only if $\delta_{\mathrm{IR}} = 0$. Under positivity, we obtain the following bound (see § D.2):

$$\|\beta_{\mathrm{OVR}} - \beta_0\|_2 \le \frac{\delta_{\mathrm{IR}}}{\sqrt{\lambda_{\min}(\Sigma_C)}},$$

where $\lambda_{\min}(\Sigma_C)$ is the smallest eigenvalue of Σ_C. Hence, the effect of violating IR depends on both its severity, measured by δ_{IR}, and positivity: small violations of IR lead to small worst-case bias when $\lambda_{\min}(\Sigma_C)$ is bounded away from zero, but can be amplified when positivity is weak.

Violations of positivity. Positivity (Assumption 2) fails when $C(X)$ determines a nonzero linear combination of X (e.g., when C is injective). For example, if $\mathbb{E}[H|X] = AX$ and $C(X)$ recovers AX, every direction in the row space of A has zero residual variance. Under additional assumptions on the conditional variances, a modified version of OVR (we call it VOVR) can identify the causal effect even under such positivity violations; see § E for details.

3.2 Smoothed OVR: A consistent algorithm

Turning the population identification result into an estimator introduces an additional difficulty: the contrast $C(X)$ is itself unknown and must be estimated from the data. We use sample splitting to ensure independence between the estimated contrast and the sample used for the subsequent residualization step. More importantly, consistency of the contrast estimation alone may not be sufficient because conditioning can be highly sensitive to small estimation errors. For example, if $|\mathcal{E}| = 2$, $C(X) = 0$ and $\widehat{C}(X) = \delta_n X$ for scalar X, then for every $\delta_n \neq 0$, $\widehat{C}(X)$ reveals X exactly and thus 'perfect' residualization yields $\widetilde{X} = X - \mathbb{E}[X|C(X), E] = 0$, so OVR fails , even though $\widehat{C}(X) \to C(X)$ as $\delta_n \to 0$.

To limit unwanted leakage of X through the estimation error in $C(X)$, we add Gaussian noise to $\widehat{C}(X)$ at the scale of its estimation uncertainty. This makes the adjustment step less sensitive

to information about X that is present only due to estimation error. Less noise is added when the variability in contrast estimation is lower. We refer to the resulting procedure, summarized in Algorithm 1, as Smoothed OVR. Under consistent nuisance estimation and suitable calibration of the smoothing noise, Theorem 2 establishes consistency of the resulting estimator. § C contains the precise assumptions and the full theorem statement; the proof is in § D.5.

Theorem 2 (Consistency, informal)**.** *Under IR (Assumption 1), positivity (Assumption 2) and additional regularity conditions (Assumption 3), Smoothed OVR (Algorithm 1) is consistent for* β_0.

4 Three special cases

We provide three special cases of the setting described in § 2.1 for which the inclusion restriction holds and discuss the positivity condition (Assumption 2) needed for identification in each.

4.1 Multivariate latent with environment-specific loadings

We now provide sufficient identifiability conditions for (a generalized version of) the motivating example from § 2.2. Let $X \in \mathbb{R}^d$, $H \in \mathbb{R}^s$, $E \in \{1, \dots, J\}$, and $\lambda_1, \dots, \lambda_J \in \mathbb{R}^s$. Assume $\mathbb{E}[U_Y|X, E] = 0$ and

$$Y := \beta_0^\top X + \lambda_E^\top H + U_Y, \qquad \text{with} \qquad \mathbb{E}[H \mid X = x, E = e] = \mathbb{E}[H \mid X = x]. \tag{11}$$

Then $\mu_e(x) = \beta_0^\top x + \lambda_e^\top \mathbb{E}[H|X = x]$. Fix any $e_0 \in \{1, \dots, J\}$ as a reference environment. Then the contrast from (9) is given by

$$C(x) = L\,\mathbb{E}[H \mid X = x], \qquad \text{where} \qquad L_{e,\cdot} := (\lambda_e - \lambda_{e_0})^\top, \qquad e \neq e_0. \tag{12}$$

The contrast only needs to determine the baseline nuisance $\lambda_{e_0}^\top \mathbb{E}[H|X]$, not the full conditional expectation $\mathbb{E}[H|X]$. Proposition 1 gives sufficient conditions for identification in this setting. Proposition 4 in § D.6 extends the result to a general factorization of the conditional nuisances.

Proposition 1 (Low-rank inclusion restriction)**.** *Suppose* (11) *holds with* $\mathbb{E}[U_Y|X, E] = 0$*, and let* L *be defined by* (12)*. If*

$$\lambda_{e_0} \in \operatorname{rowspan}(L) \qquad \text{and} \qquad \mathbb{E}[\operatorname{Var}(X \mid \mathbb{E}[H \mid X], E)] \succ 0,$$

then Assumptions 1 and 2 hold and $\beta_{\mathrm{OVR}} = \beta_0$.

4.2 Existence of at least one unconfounded environment

A simple sufficient condition arises when one environment has no X-dependent bias. In that case, a contrast with this environment directly reveals the nuisance in the remaining environments.

Proposition 2. *Suppose some* $e_\star \in \mathcal{E}$ *satisfies* $q_{e_\star}(X) = \kappa_\star$ *almost surely for a constant* $\kappa_\star$*. If additionally* $\Sigma_C \succ 0$ *(Assumption 2), then* $\beta_{\mathrm{OVR}} = \beta_0$.

Notably, knowledge of which environment is unconfounded is not required (so $e_\star$ can be unknown). Further, the assumption requires only that $q_{e_\star}(X)$ be constant; it does not require $H \perp\!\!\!\perp X \mid E = e_\star$.

4.3 Recoverable confounding state

In this setting, a latent state Z_C (see Fig. 4) that can be recovered from the treatment X suffices to adjust for the confounding within each environment.

Proposition 3 (Recoverable confounding state)**.** *Suppose* $Y := \beta_0^\top X + \psi(H, E) + U_Y$ *with* $\mathbb{E}[U_Y|X, E] = 0$*. Suppose there are* $Z_C \in \mathbb{R}^r$ *and a measurable function* F *such that* $Z_C = F(X)$ *almost surely and* $X \perp\!\!\!\perp H|(Z_C, E)$*. For each* $e \in \mathcal{E}$*, define* $\rho_e(z) := \mathbb{E}[\psi(H, e)|Z_C = z, E = e]$ *and* $\Delta(z) := (\rho_e(z) - \rho_{e_0}(z))_{e \neq e_0}$*. If* Δ *is injective on the support of* Z_C *and* $\Sigma_{Z_C} := \mathbb{E}[\operatorname{Var}(X|Z_C, E)] \succ 0$*, then Assumptions 1 and 2 hold,* $\Sigma_C = \Sigma_{Z_C}$*, and* $\beta_{\mathrm{OVR}} = \beta_0$.

For a concrete example, consider an SCM with the graph in Fig. 4. Here, the scalar state $Z_C > 0$ is the treatment vector's Euclidean length: $Z_C = \|X\|$ almost surely. Write $X = Z_C V \in \mathbb{R}^d$, where $V \in \mathbb{R}^d$ is its direction and $\|V\| = 1$ almost surely. Let the (scalar) confounder H and outcome Y satisfy $H := Z_C^2 + b_E Z_C + U_H$, and $Y := \beta_0^\top X + H + d_E + U_Y$, where U_H and U_Y are mutually independent, centered scalar noises, independent of (Z_C, V, E); and $(b_e, d_e)_{e \in \mathcal{E}}$ are environment-specific constants.

For fixed length Z_C and environment E, the direction V provides no additional information about H, so $X \perp\!\!\!\perp H | (Z_C, E)$. Here, $\rho_e(z) = z^2 + b_e z + d_e$ and $q_e(X) = \rho_e(Z_C)$. Hence, $C(X) = \Delta(Z_C)$, where $\Delta(z) = \big((b_e - b_{e_0})z + (d_e - d_{e_0})\big)_{e \neq e_0}$. If $b_e \neq b_{e_0}$ for at least one environment, then Δ is injective. If additionally $\mathbb{E}[Z_C^2 \operatorname{Var}(V | Z_C, E)] \succ 0$, Proposition 3 implies $\beta_{\mathrm{OVR}} = \beta_0$. To apply OVR, we use $C(X)$ and E as adjustment variables; we need not specify that $Z_C = \|X\|$.

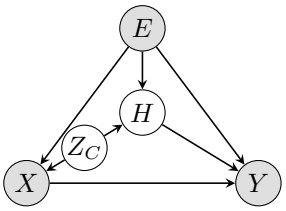


Figure 4: Setting with recoverable latent-state Z_C.

5 Discussion in the context of related work

Several methods identify coefficients of linear structural equation models using changes in additive noise distributions across environments, while keeping the structural coefficients fixed. Causal Dantzig (Rothenhäusler et al., 2019) requires $\mathbb{E}[X(Y - \beta_0^\top X) | E = e]$ to be the same in every environment and estimates β_0 by differencing these moment equations. KL regression (Roquero Gimenez & Zou, 2020) fits a common causal slope to the covariance matrices of (X, Y) across environments; it allows the noise covariance of X and noise variance of Y to change while keeping the confounder distribution and its effects fixed. BackShift (Rothenhäusler et al., 2015) estimates the structural coefficients among all observed variables (allowing feedback loops) from covariance differences. It assumes an invariant baseline noise distribution and additive shifts that are uncorrelated across variables and with the baseline noise. Our setting allows the effect of H on Y to change across environments while β_0 remains fixed. Such changes can violate Causal Dantzig's moment invariance and the invariant confounding structure assumed by KL regression and BackShift, so their identification guarantees do not cover our setting in general. OVR forms contrasts of the conditional expectations $\mu_e(x)$ at fixed x and adjusts for these contrasts under inclusion restriction and positivity.

Perhaps the most closely related result to our work is Thm. 3.4 of Kivva et al. (2025), which identifies a scalar effect when one latent-outcome loading changes across two environments, using a linear independent-noise model and higher-moment or non-Gaussianity conditions. In their loading-shift setting, IR holds because the contrast is a nonzero multiple of $\mathbb{E}[H|X]$; positivity, however, fails if this conditional expectation is injective in X. Our identification result allows vector treatments and nonlinear conditional nuisances under inclusion restriction and positivity. It requires finite second moments but no parametric noise distribution or higher-order-moment identifying equations.

Unlike invariance-based and adjustment-set methods (Peters et al., 2016; Arjovsky et al., 2019; De Bartolomeis et al., 2025), OVR does not seek an invariant predictor nor an observed adjustment set.

Environment IV (Newey et al., 1999), anchor regression (Rothenhäusler et al., 2021), and control-function approaches (Gnecco et al., 2026) exploit environment-induced variation in the treatment X, with method-specific exclusion or stability conditions OVR instead leverages variation in the nuisance $q_e(x)$: the inclusion restriction requires $C(X)$ to determine $q_{e_0}(X)$, and positivity requires residual variation in X given $(C(X), E)$. Under these assumptions, OVR can identify β_0 without an $E \to X$ shift. Conversely, invariant confounding terms cancel from the contrast and thus cannot be handled by OVR, but a valid environment instrument may still identify β_0 in such a setting.

Proxy-variable methods use additional observed variables that carry information about H, together with proxy-exclusion restrictions and bridge or completeness conditions (Miao et al., 2018; Tchetgen Tchetgen et al., 2024). OVR uses a different object: $C(X)$ is generated from environment-specific conditional expectations and is a deterministic function of X, not an additional measurement of H. IR requires $C(X)$ to determine $q_{e_0}(X)$; it does not require recovering H or solving a bridge equation. The approaches are therefore complementary. Proxy methods do not require environment variation when suitable proxies are observed, whereas OVR does not require a measured proxy but requires informative outcome-mechanism shifts and positivity. § A discusses further related work.

6 Experiments

We use Smoothed OVR (Algorithm 1) with smoothing scale $\tau = 1$. For each observation, we average predictions at $M = 8$ perturbed contrast inputs, using four independent Gaussian noise

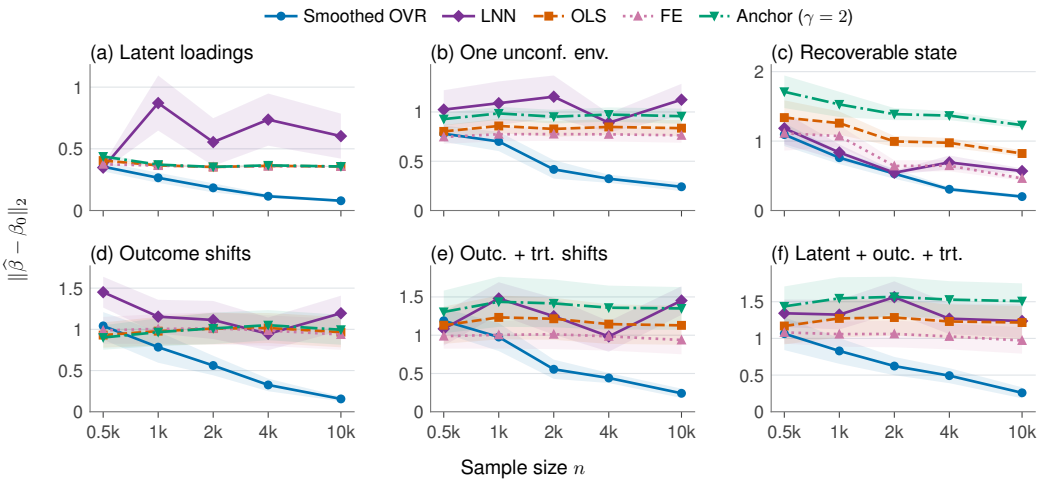


Figure 5: Coefficient error at $n \in \{0.5, 1, 2, 4, 10\}$k across six synthetic settings. Lines show means and bands $\pm$ one Monte Carlo s.e. over ten paired replications; panels use separate linear error scales. As n increases, Smoothed OVR has the lowest mean ℓ_2 estimation error in all six settings.

draws and their negatives. § F gives implementation and experimental details. § F.4 compares Direct (based on simple plug-in estimation) and Smoothed OVR, thus isolating the effect of smoothing.

6.1 Synthetic study

We compare Smoothed OVR (Algorithm 1) with linear regression of Y on X (OLS), linear environment fixed-effects regression (FE), a jointly trained linear-plus-neural-network model (LNN), which fits $Y \approx \beta^\top X + f_\theta(X, E)$ by squared error (§ F.2), and anchor regression (Rothenhäusler et al., 2021) with $\gamma = 2$. The estimator of Kivva et al. (2025) assumes a scalar treatment and a linear independent-noise SCM and is therefore not directly applicable to our setting, where treatments are multivariate and the confounding mechanisms need not be linear.

We consider several different data generating processes. Fig. 5 reports the means $\|\widehat{\beta} - \beta_0\|_2$ and standard errors for different estimators over ten replications, using the same generated observations for all methods within each replication, at sample sizes $n \in \{500, 1000, 2000, 4000, 10000\}$. Panels (a)–(c) correspond to the three sufficient settings in § 4. Panels (d)–(f) do not impose IR and progressively add shifts in the outcome mechanism, the treatment distribution, and the latent-confounder distribution. The precise SCMs and implementation details are specified in § F.

(a) Latent loadings (§ 4.1). Two binary confounders H_1, H_2 affect disjoint pairs of treatment components: H_1 affects (X_1, X_2) and H_2 affects (X_3, X_4). The environment E changes only their coefficients in the outcome equation. The loading-difference matrix L has full column rank, so $\mathbb{E}[(H_1, H_2)^\top | X]$ can be recovered from $C(X)$. Variation within each pair ensures positivity, satisfying the conditions of Proposition 1.

(b) One unconfounded environment (§ 4.2). The environment E induces additive shifts in the equations for H and X, with X depending nonlinearly on H. One environment (it is unknown to the method which one) is unconfounded ($q_{e_\star} = 0$). The contrast with this environment reveals the reference nuisance as in Proposition 2, while Gaussian treatment noise gives common support.

(c) Recoverable confounding state (§ 4.3). We use the special case after Proposition 3 with $Z_C = \|X\|$: within each environment, $q_e(X)$ depends only on this length. After adjustment, variation in the direction of X leaves positive residual variance in every nonzero linear combination of X, ensuring positivity.

(d) Outcome shifts. Here $E \perp\!\!\!\perp (H, X)$, and Y is the only child of E, as in Fig. 2 without the dashed edges. A non-linear function maps a scalar confounder H and independent noise to $X \in \mathbb{R}^4$. The contribution of H to Y is represented by a separate neural network in each environment, with

Table 1: Mean ℓ_2 estimation error over ten training/reference splits, with Monte Carlo standard errors in parentheses and the lowest means in bold.

Smoothed OVR	OLS	FE	Anchor ($\gamma = 2$)	LNN
.019 (.001)	.129 (.001)	.104 (.001)	.171 (.001)	.095 (.011)

weights drawn independently across environments. We do not impose the structure used in § 4 to ensure IR, so $q_{e_0}(X)$ is generally not perfectly recoverable from $C(X)$.

(e) Outcome and treatment shifts. Starting from (d), we add an environment-specific mean shift to X (that is, $E \to X$ in Fig. 2). Thus $P(H|X, E)$ can now depend on E, even though $E \perp\!\!\!\perp H$ still holds. As in (d), we impose neither IR nor positivity.

(f) Outcome, treatment, and confounder shifts. Starting from (e), we also add an environment-specific mean shift to H (that is, $E \to H$ in Fig. 2). Thus E now directly affects H, X, and Y. As in (d) and (e), we impose neither IR nor positivity.

Fig. 5 shows that the mean ℓ_2 estimation error of Smoothed OVR decreases with n in all six settings. Smoothed OVR is competitive at smaller sample sizes and has the lowest mean error in every setting for sample sizes $n > 2{,}000$. Panels (d)–(f) impose neither IR nor positivity, so identification is not guaranteed by construction. The contrasts may nevertheless remove much of the confounding; the bound in § 3.1 shows that population bias can remain small when violations of IR are small and sufficient treatment variation remains after adjustment.

6.2 Real-data benchmark

We use KuaiRand-Pure (Gao et al., 2022), a dataset of user-video interactions from the Kuaishou short-video platform. Videos are selected either by the platform's usual recommender system (standard policy) or randomly from a candidate pool (random policy). We use these two assignment policies as the environments E. We retain users and videos observed under both policies and sample equal policy counts within each user. The resulting cohort has 339,954 interactions, 20,669 users, and 5,613 videos. The standardized treatment vector $X \in \mathbb{R}^6$ contains log video duration, log upload age, log aspect ratio, indicators for long-import and web uploads, and tag count. The outcome is $Y = \log(1 + \min\{\text{watch seconds}, 300\})$.

For each of ten repetitions, we split interactions separately for each user and policy. The resulting two training environments contain 80,079 standard and 80,079 random interactions; all methods use these same 160,158 observations. In each repetition, we fit a separate linear regression of Y on X using the remaining 89,898 random-policy interactions. Its fitted coefficients serve as the reference for evaluation. No observations appears in both the training and reference sets. The same 20,669 users appear in both samples to keep the target population fixed. The random policy breaks assignment based on the recommender's user state, but does not independently randomize the six video attributes or eliminate unmeasured item characteristics. We therefore evaluate distance to the linear-regression coefficients under random video assignment; these are not known structural coefficients. Crucially, none of the methods are told which of the two training environments corresponds to the reference random policy.

We evaluate all methods by their mean ℓ_2 estimation error to the reference coefficients, see Table 1 for the results. Smoothed OVR achieves the lowest mean ℓ_2 estimation error over ten training/reference splits (.019), compared with .095 for LNN, the closest baseline.

7 Conclusion and future work

Outstrumental Variable Regression complements methods that require environment-induced treatment shifts such as IV by turning outcome heterogeneity into an adjustment signal. We establish identifiability under the inclusion restriction and positivity, and consistency of the smoothed estimator under additional regularity conditions. Future work includes extensions to non-linear effects, e.g. via feature maps, and to continuous environments, where the contrast becomes function-valued.

Reproducibility statement

Proofs of all theoretical results are provided in Appendix D. Implementation details, data-generating processes, hyperparameters, and additional experiments are given in Appendix F. Code to reproduce the experiments is provided in under https://anonymous.4open.science/r/ovr-B865.

AI use statement

Generative AI tools were used to assist with literature search and related-work discovery, research ideation and methodological development, code development, manuscript drafting and editing, and the development and checking of mathematical arguments and proofs. All AI-assisted research ideas, mathematical claims, proofs, code, experimental results, and citations were reviewed and verified by the authors. The authors take full responsibility for the final content of the paper.

References

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, 105(490):493–505, 2010. 14

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2009. 1

Joshua D Angrist, Guido W Imbens, and Donald B Rubin. Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455, 1996. 1

Martin Arjovsky, Léon Bottou, Ishaan Gulrajani, and David Lopez-Paz. Invariant risk minimization. *arXiv preprint arXiv:1907.02893*, 2019. 7, 13

Susan Athey, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, 116(536):1716–1730, 2021. 14

Stephan Bongers, Patrick Forré, Jonas Peters, and Joris M. Mooij. Foundations of Structural Causal Models with Cycles and Latent Variables. *The Annals of Statistics*, 49(5), October 2021. 3

David Card and Alan B Krueger. Minimum wages and employment: A case study of the fast food industry in new jersey and pennsylvania, 1993. 14

Bénédicte Colnet, Imke Mayer, Guanhua Chen, Awa Dieng, Ruohong Li, Gaël Varoquaux, Jean-Philippe Vert, Julie Josse, and Shu Yang. Causal inference methods for combining randomized trials and observational studies: a review. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 39(1):165, 2024. 14

Alexander D'Amour. On multi-cause approaches to causal inference with unobserved counfounding: Two cautionary failure cases and a promising alternative. In *The 22nd international conference on artificial intelligence and statistics*, pp. 3478–3486. PMLR, 2019. 14

Piersilvio De Bartolomeis, Julia Kostin, Javier Abad, Yixin Wang, and Fanny Yang. Doubly robust identification of treatment effects from multiple environments. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=9vTAkJ9Tik. 7, 13

Ronald A Fisher. *The Design of Experiments*. Oliver and Boyd, Edinburgh, 1935. 1

Chongming Gao, Shijun Li, Yuan Zhang, Jiawei Chen, Biao Li, Wenqiang Lei, Peng Jiang, and Xiangnan He. KuaiRand: An unbiased sequential recommendation dataset with randomly exposed videos. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, pp. 3953–3957, 2022. doi: 10.1145/3511808.3557624. URL https://doi.org/10.1145/3511808.3557624. 9

Nicola Gnecco, Jonas Peters, Sebastian Engelke, and Niklas Pfister. Boosted control functions: Distribution generalization and invariance in confounded models. *Journal of Machine Learning Research*, 27(46):1–57, 2026. URL https://www.jmlr.org/papers/v27/24-2207.html. 2, 7, 14

Tobias Hatt, Jeroen Berrevoets, Alicia Curth, Stefan Feuerriegel, and Mihaela van der Schaar. Combining observational and randomized data for estimating heterogeneous treatment effects. *arXiv preprint arXiv:2202.12891*, 2022. 14

Aapo Hyvarinen, Hiroaki Sasaki, and Richard Turner. Nonlinear ICA using auxiliary variables and generalized contrastive learning. In *International Conference on Artificial Intelligence and Statistics*, pp. 859–868. PMLR, 2019. 13

Manuel Iglesias-Alonso, Felix Schur, Julius von Kügelgen, and Jonas Peters. Transferring causal effects using proxies. *Advances in Neural Information Processing Systems*, 38:111360–111393, 2026. 14

Nathan Kallus, Aahlad Manas Puli, and Uri Shalit. Removing hidden confounding by experimental grounding. *Advances in neural information processing systems*, 31, 2018. 14

Rickard K. A. Karlsson and Jesse H. Krijthe. Detecting hidden confounding in observational data using multiple environments. *Advances in Neural Information Processing Systems*, 36:44280–44309, 2023. URL https://papers.nips.cc/paper_files/paper/2023/hash/89e541b817ea043a971840a926e12b37-Abstract-Conference.html. 13

Ilyes Khemakhem, Diederik Kingma, Ricardo Monti, and Aapo Hyvarinen. Variational autoencoders and nonlinear ica: A unifying framework. In *International conference on artificial intelligence and statistics*, pp. 2207–2217. PMLR, 2020. 13

Yaroslav Kivva, Sina Akbari, Saber Salehkaleybar, and Negar Kiyavash. Causal effect identification in heterogeneous environments from higher-order moments. In *Proceedings of the Forty-first Conference on Uncertainty in Artificial Intelligence*, volume 286 of *Proceedings of Machine Learning Research*, pp. 2223–2254. PMLR, 2025. URL https://proceedings.mlr.press/v286/kivva25a.html. 7, 8, 13

Marc Lipsitch, Eric Tchetgen Tchetgen, and Ted Cohen. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology (Cambridge, Mass.)*, 21(3):383, 2010. 14

Sarah Mameche, Jilles Vreeken, and David Kaltenpoth. Identifying confounding from causal mechanism shifts. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 4897–4905. PMLR, 2024. URL https://proceedings.mlr.press/v238/mameche24a.html. 13

Wang Miao, Zhi Geng, and Eric J. Tchetgen Tchetgen. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993, 2018. doi: 10.1093/biomet/asy038. 7, 14

Whitney K Newey, James L Powell, and Francis Vella. Nonparametric estimation of triangular simultaneous equations models. *Econometrica*, 67(3):565–603, 1999. 7, 14

Chan Park, David B Richardson, and Eric J Tchetgen Tchetgen. Single proxy control. *Biometrics*, 80(2):ujae027, 2024. 14

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2nd edition, 2009. 1, 3

Jonas Peters, Peter Bühlmann, and Nicolai Meinshausen. Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(5):947–1012, 2016. 7, 13

Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. The MIT press, 2017. 1

Jaime Roquero Gimenez and James Zou. Identifying invariant factors across multiple environments with KL regression. *arXiv preprint arXiv:2002.08341*, 2020. 7, 13

Evan TR Rosenman, Guillaume Basse, Art B Owen, and Mike Baiocchi. Combining observational and experimental datasets using shrinkage estimators. *Biometrics*, 79(4):2961–2973, 2023. 14

Jonathan Roth, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of econometrics*, 235 (2):2218–2244, 2023. 14

Dominik Rothenhäusler, Christina Heinze, Jonas Peters, and Nicolai Meinshausen. BackShift: Learning causal cyclic graphs from unknown shift interventions. In *Advances in Neural Information Processing Systems 28*. Curran Associates, Inc., 2015. 7, 13

Dominik Rothenhäusler, Peter Bühlmann, and Nicolai Meinshausen. Causal Dantzig: Fast inference in linear structural equation models with hidden variables under additive interventions. *The Annals of Statistics*, 47(3):1688–1722, 2019. doi: 10.1214/18-AOS1732. 2, 7, 13

Dominik Rothenhäusler, Nicolai Meinshausen, Peter Bühlmann, and Jonas Peters. Anchor regression: Heterogeneous data meet causality. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 83(2):215–246, 2021. 2, 7, 8, 14

Felix Schur, Pio Blieske, and Jonas Peters. Decor: deconfounding time series with robust regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 88(3):799–818, 2026a. 14

Felix Schur, Niklas Pfister, Peng Ding, Sach Mukherjee, and Jonas Peters. Many experiments, few repetitions, unpaired data, and sparse effects: Is causal inference possible? *arXiv preprint arXiv:2601.15254*, 2026b. 2

Eric Tchetgen Tchetgen. The control outcome calibration approach for causal inference with unobserved confounding. *American journal of epidemiology*, 179(5):633–640, 2014. 14

Eric J. Tchetgen Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao. An introduction to proximal causal inference. *Statistical Science*, 39(3):375–390, 2024. doi: 10.1214/23-STS911. 7, 14

Nikolaj Thams, Rikke Søndergaard, Sebastian Weichwald, and Jonas Peters. Identifying causal effects using instrumental time series: Nuisance iv and correcting for the past. *Journal of Machine Learning Research*, 25(302):1–51, 2024. 13

Silvana Tiedemann, Jorge Sanchez Canales, Felix Schur, Raffaele Sgarlato, Lion Hirth, Oliver Ruhnau, and Jonas Peters. Identifying elasticities in autocorrelated time series using causal graphs. *arXiv preprint arXiv:2409.15530*, 2024. 13

Burak Varici, Emre Acartürk, Karthikeyan Shanmugam, Abhishek Kumar, and Ali Tajer. Score-based causal representation learning: Linear and general transformations. *Journal of Machine Learning Research*, 26(112):1–90, 2025. 13

Victor Veitch, Dhanya Sridhar, and David Blei. Adapting text embeddings for causal inference. In *Conference on uncertainty in artificial intelligence*, pp. 919–928. PMLR, 2020. 14

Silvan Vollmer, Niklas Pfister, and Sebastian Weichwald. Identifying causal effects using a single proxy variable. *arXiv preprint arXiv:2604.09135*, 2026. 14

Julius von Kügelgen, Michel Besserve, Liang Wendong, Luigi Gresele, Armin Kekić, Elias Bareinboim, David Blei, and Bernhard Schölkopf. Nonparametric identifiability of causal representations from unknown interventions. *Advances in Neural Information Processing Systems*, 36: 48603–48638, 2023. 13

Yixin Wang and David M Blei. The blessings of multiple causes. *Journal of the American Statistical Association*, 114(528):1574–1596, 2019. 14

Liang Wendong, Armin Kekić, Julius von Kügelgen, Simon Buchholz, Michel Besserve, Luigi Gresele, and Bernhard Schölkopf. Causal component analysis. *Advances in Neural Information Processing Systems*, 36:32481–32520, 2023. 13

Yiqing Xu. Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76, 2017. 14

Shu Yang and Peng Ding. Combining multiple observational data sources to estimate causal effects. *Journal of the American Statistical Association*, 115(531):1540–1554, 2020. 14

Shu Yang, Chenyin Gao, Donglin Zeng, and Xiaofei Wang. Elastic integrative analysis of randomised trial and real-world data for treatment heterogeneity estimation. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(3):575–596, 2023. 14

A Extended Related Work

A.1 Multi-environment identifiability and invariance

Causal Dantzig identifies a linear effect from cross-environment changes in equations involving $\mathbb{E}[XY]$ and $\mathbb{E}[XX^\top]$ under additive interventions (Rothenhäusler et al., 2019), while KL regression fits a common linear-SEM coefficient to environment-specific joint covariances under shift interventions (Roquero Gimenez & Zou, 2020). BackShift targets the entire connectivity matrix of a linear, potentially cyclic equilibrium SEM by jointly diagonalizing covariance differences (Rothenhäusler et al., 2015). It permits latent confounding through a non-diagonal noise covariance, but requires both that this covariance and the structural matrix remain invariant and that the unknown intervention shifts be uncorrelated with the structural noise and have diagonal covariance. A changing latent-outcome loading—including our motivating outcome-only mechanism shift—violates these invariance conditions; moreover, BackShift's generic identifiability result needs at least three environments, whereas Outstrumental Variable Regression can identify with two. Thus covariance-shift methods exploit changing linear second moments under invariant structural systems, while Outstrumental Variable Regression permits the conditional confounding nuisance itself to vary and can identify with $P(X \mid E)$ unchanged.

The closest outcome-side result is Theorem 3.4 of Kivva et al. (2025), which also allows an unchanged treatment distribution when a latent-outcome loading changes, but assumes a scalar linear independent-noise graph and uses higher-moment or non-Gaussianity conditions. Outstrumental Variable Regression differences the full conditional expectations, permitting vector treatments and nonlinear conditional nuisances under an invariant partially linear slope, IR, and positivity.

RAMEN searches for an observed adjustment set through conditional-mean invariance and requires the corresponding parents to be observed (De Bartolomeis et al., 2025); Outstrumental Variable Regression instead uses outcome non-invariance to construct an adjustment score. Work that leverages mechanism shifts to detect hidden confounding, rather than identify its average causal effect, addresses a different target (Karlsson & Krijthe, 2023; Mameche et al., 2024). Invariant causal prediction tests observed covariate subsets for an environment-invariant conditional outcome law and targets causal parents (Peters et al., 2016), whereas invariant risk minimization learns representations supporting a common risk-optimal predictor for out-of-distribution prediction (Arjovsky et al., 2019). Outstrumental Variable Regression uses the opposite signal: structured outcome non-invariance reveals the conditional nuisance while the causal effect remains invariant.

The broader idea that sufficiently rich environmental variation can reveal otherwise latent structure also appears in nonlinear independent component analysis with auxiliary variables (Hyvarinen et al., 2019; Khemakhem et al., 2020) and interventional causal representation learning (von Kügelgen et al., 2023; Wendong et al., 2023; Varici et al., 2025). In these settings, identifiability likewise relies on sufficiently diverse changes across environments, although the objects being contrasted and the resulting identification problems are different.

A.2 Instruments and control functions

Environment IV methods use E as an instrument, as depicted in Fig. 1, and extend to time series (Thams et al., 2024; Tiedemann et al., 2024), while classical control functions construct a control

from a treatment first stage under triangular-model exclusion and stability restrictions (Newey et al., 1999). Anchor regression approaches an IV/2SLS-type estimand as its robustness penalty grows; that limit identifies the structural slope only under the corresponding instrument-validity conditions (Rothenhäusler et al., 2021). Under a simultaneous-equation shift model and sufficient heterogeneity, boosted control functions instead target an invariant worst-case predictor that need not equal the structural response (Gnecco et al., 2026). Outstrumental Variable Regression does not use E as an instrument: it allows a direct environment-dependent outcome nuisance and can operate when $E \perp\!\!\!\perp X$. Conversely, valid instruments and covariance-based methods can identify affine regimes in which Outstrumental Variable Regression fails.

A.3 Proxy-based deconfounding and semiparametric estimation

The deconfounder infers a substitute confounder from a multi-cause assignment model (Wang & Blei, 2019), although multiple causes alone do not generally identify causal effects (D'Amour, 2019); causally supervised text embeddings instead compress an observed variable assumed sufficient for adjustment (Veitch et al., 2020). Proximal methods begin with measured treatment- and outcome-side proxies and impose proxy exclusions, bridge existence, and completeness conditions (Miao et al., 2018; Tchetgen Tchetgen et al., 2024). Proxy information can also enable the transfer of causal effects across domains under unobserved confounding (Iglesias-Alonso et al., 2026). Single-proxy identifiability instead relies on a negative-control outcome or a known proxy mechanism plus completeness (Park et al., 2024; Vollmer et al., 2026). The contrast $C(X)$ has different auxiliary structure: it is generated from environment-specific conditional expectations and is a deterministic function of X, rather than a measured proxy for H or a proximal bridge. IR requires it to balance the conditional nuisance, not to recover the latent confounder itself. Because the score is estimated, the slope is fitted on a separate sample split; § 3.2 and § C analyze the resulting estimator.

A.4 Alternative structure for unobserved confounding

A different line of work obtains causal identification by exploiting structure in the unobserved heterogeneity rather than exclusion restrictions or measured proxies. In panel settings, synthetic-control and interactive-fixed-effect methods leverage cross-unit and temporal structure that can be represented through latent factor models, while matrix-completion approaches impose related low-rank structure on untreated potential outcomes (Abadie et al., 2010; Xu, 2017; Athey et al., 2021). For time-series data, Schur et al. (2026a) instead exploit spectral sparsity of an unobserved confounding process and reduce causal effect estimation to robust regression in the frequency domain. These approaches are complementary to OVR: they rely on structural restrictions across units or time, whereas OVR exploits informative variation of the outcome mechanism across environments.

A.5 Differencing across data sources to expose confounding

The elementary step behind OVR, differencing quantities across which the causal component is invariant so that only the confounding component remains, also underlies several established approaches. Difference-in-differences compares outcome changes between treated and untreated groups under parallel trends, that is, an additive, covariate-invariant confounding shift (Card & Krueger, 1993; Roth et al., 2023). Negative controls and control-outcome calibration use an auxiliary outcome known to be unaffected by the treatment, whose observed association with the treatment calibrates the confounding bias in the primary outcome (Lipsitch et al., 2010; Tchetgen Tchetgen, 2014); the single-proxy methods discussed above build on the same idea. Experimental grounding fits a flexible effect model on confounded observational data and corrects it by a parametric function of the difference to a small unconfounded trial (Kallus et al., 2018). More generally, methods that combine randomized and observational data assume that one source is unconfounded, or observes the full confounder set, and use the other source for efficiency or extrapolation (Yang & Ding, 2020; Hatt et al., 2022; Rosenman et al., 2023; Yang et al., 2023); see Colnet et al. (2024) for a review.

OVR differs from this line of work in two respects. First, no environment needs to be unconfounded, and no environment needs to be known to be clean. Proposition 2 covers the special case in which one environment has a covariate-independent nuisance; there, OVR recovers the logic of experimental grounding without requiring the clean environment to be labeled. Proposition 1 and Proposition 3

instead identify β_0 from environments that are all confounded but load differently on a common confounding score, a setting in which no unconfounded source is supplied. Second, the contrast is not added to a confounded estimate as a correction; it is used as a conditioning variable, and identification comes from the treatment variation that remains within its level sets (Assumption 2). If $C(X)$ is constant, OVR reduces to environment fixed-effects regression.

B Estimation Details

Overlap and contrast estimability. Each coordinate of $C(x)$ compares conditional expectations from two environments at the same x. The environment overlap condition (2) identifies every μ_e, and hence C, P_X-almost surely; a uniform lower bound on $P(E = e \mid X)$ is typically imposed when deriving first-stage rates for the contrast-convergence requirement in Assumption 3(a). Overlap is distinct from the residual-variation positivity condition in Assumption 2.

Direct OVR (DOVR; Algorithm 2) implements the residual regression in (10) using estimated contrasts without smoothing. We abbreviate Smoothed OVR (Algorithm 1) as SOVR.

Algorithm 2 Direct Outstrumental Variable Regression

Input: Data $\{(X_i, Y_i, E_i)\}_{i=1}^n$, reference e_0, folds K.

1. Split the data into $\mathcal{I}_C$ and $\mathcal{I}_\beta$. On $\mathcal{I}_C$, estimate $\widehat{\mu}_e$ and set
$$\widehat{C}(x) \coloneqq \big(\widehat{\mu}_e(x) - \widehat{\mu}_{e_0}(x)\big)_{e \neq e_0}.$$
2. Cross-fit regressions of X and Y on $(\widehat{C}(X), E)$ within $\mathcal{I}_\beta$, yielding held-out residuals $(\widehat{X}_i, \widehat{Y}_i)$.
3. Return
$$\widehat{\beta}_{\text{direct}} = \left(\sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{X}_i^\top\right)^{-1} \sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{Y}_i.$$

The direct estimator first fits $\mu_e(x) = f(x) + c_e(x)$ jointly across environments, with a shared component f and contrast heads satisfying $c_{e_0} \equiv 0$. Joint fitting helps prevent the invariant signal from leaking into separately estimated contrasts. The smoothed procedure in Algorithm 1 instead uses two contrast fits and noisy-input residualization; Theorem 3 establishes its consistency.

B.1 Diagnostics

Let $\widehat{X}_i$ be the held-out treatment residuals, using averaged noisy predictions for Smoothed OVR, and set $\widehat{\Sigma} = |\mathcal{I}_\beta|^{-1} \sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{X}_i^\top$.. With $\widehat{\Sigma}_E$ the covariance after residualizing X only on E, we report the positivity diagnostic $\widehat{\gamma} = \lambda_{\min}(\widehat{\Sigma}_E^{-1/2} \widehat{\Sigma} \widehat{\Sigma}_E^{-1/2})$, which estimates the smallest fraction of within-environment linear treatment variation retained after adjustment. Small values warn that positivity is nearly lost.

C Consistency of Smoothed OVR

Sampling and notation. We observe an i.i.d. sample from P as in § 2.1 and use the decomposition $Y = \beta_0^\top X + q_E(X) + W$ from § 2.3. Under Assumption 1, $q_E(X) = \varphi_E(C(X))$ with $\varphi_e(c) \coloneqq \varphi(c) + c_e$ and $c_{e_0} \coloneqq 0$. We analyze Algorithm 1 with splits drawn independently of the observations, $|\mathcal{I}_C| = n_C \to \infty$, $|\mathcal{I}_\beta| = n \to \infty$, and a fixed number K of folds with $|I_k| \to \infty$. Environment-stratified splits are also covered if every split and fold has limiting environment proportions $\Pr(E = e)$; the laws of large numbers below are then applied within each environment. Write $\widehat{C}_n \coloneqq \overline{C}$, $V_{n,k} \coloneqq \tau^2 \widehat{\Omega}_k$ for fixed τ, and $f_{n,k,Z} \coloneqq \widehat{f}_{Z,k}$ for $Z \in \{X, Y\}$.

Let $\mathcal{T}_{n,k}$ contain the first-stage fits, the training observations outside I_k, and the nuisance-fitting randomness. For a fresh observation independent of $\mathcal{T}_{n,k}$, let $U_{n,k} = \widehat{C}_n(X) + \xi$, where $\xi \mid \mathcal{T}_{n,k} \sim$

$N(0, V_{n,k})$ is independent of the observation, and define

$$m_{n,k,Z}(u, e) = \mathbb{E}[Z \mid U_{n,k} = u, E = e, \mathcal{T}_{n,k}], \qquad Z \in \{X, Y, q_E(X)\}.$$

The analysis covers any perturbation scheme in which each held-out row uses M perturbations $\xi_{i1}, \ldots, \xi_{iM}$ with marginals $N(0, V_{n,k})$, shared by the X and Y predictions, independent across rows and of the held-out observations given $\mathcal{T}_{n,k}$; the antithetic pairs of Algorithm 1 are a special case. Let $\widehat{G}_n = n^{-1} \sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{X}_i^\top$ and $\widehat{\beta}_{\mathrm{SOVR}} = (\widehat{G}_n + \lambda_n I)^{-1} n^{-1} \sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{Y}_i$, where $\lambda_n \geq 0$ tends to zero; $\lambda_n = 0$ gives the estimator of Algorithm 1, set arbitrarily if $\widehat{G}_n$ is singular.

Assumption 3 (Smoothed consistency regularity)**.** *For every fold k:*

(a) The contrast error satisfies $\rho_n = \|\widehat{C}_n - C\|_{L^2(P_X)} \to_P 0$, and $\mathrm{tr}(V_{n,k}) \to_P 0$.

(b) With $\delta_n(X) = \widehat{C}_n(X) - C(X)$ and $\zeta_{n,k} = \delta_n(X) - \mathbb{E}[\delta_n(X) \mid C(X), E, \mathcal{T}_{n,k}]$, the leakage satisfies

$$\zeta_{n,k} \in \mathrm{range}(V_{n,k}) \text{ a.s.}, \qquad \kappa_{n,k} = \mathbb{E}[\zeta_{n,k}^\top V_{n,k}^\dagger \zeta_{n,k} \mid \mathcal{T}_{n,k}] = O_P(1).$$

Here $V^\dagger$ is the Moore–Penrose inverse; the range condition excludes unperturbed leakage when V is singular.

(c) For $Z = X, Y$, the noisy-input regression error satisfies

$$\mathbb{E}\big[\|f_{n,k,Z}(U_{n,k}, E) - m_{n,k,Z}(U_{n,k}, E)\|^2 \mid \mathcal{T}_{n,k}\big] = o_P(1).$$

Part (b) involves only the component of the contrast error that carries information beyond (C, E). § C.1 gives a sufficient condition for parts (a) and (b).

Theorem 3 (Consistency of Smoothed OVR)**.** *Under the sampling scheme above and Assumptions 1 to 3, for every fixed M with the stated Gaussian marginals,*

$$\Pr(\widehat{G}_n \succ 0) \longrightarrow 1, \qquad \|\widehat{G}_n^{-1}\| = O_P(1), \qquad \widehat{\beta}_{\mathrm{SOVR}} \xrightarrow{p} \beta_0.$$

The proof in § D.5 separates recovery of the confounding term from preservation of treatment variation.

C.1 Sufficient condition for disagreement calibration

Correctly specified regression on a fixed set of features gives a concrete sufficient condition for Assumption 3(a) and (b). Let $d_C := \dim C = |\mathcal{E}| - 1$ and let $\phi(X) \in \mathbb{R}^p$, with $p \geq d_C$, satisfy $\mathbb{E}\|\phi(X)\|^2 < \infty$ and $\Gamma = \mathrm{Cov}(\phi(X)) \succ 0$. Suppose

$$C(x) = B_0 \phi(x), \qquad \widehat{C}^s(x) = \widehat{B}^s \phi(x), \quad s \in \{A, B\},$$

where $B_0, \widehat{B}^A, \widehat{B}^B \in \mathbb{R}^{d_C \times p}$. Assume the two fits, based on disjoint halves of the contrast sample, satisfy

$$\sqrt{n_C}\big(\widehat{B}^A - B_0, \widehat{B}^B - B_0\big) \Rightarrow (G_A, G_B),$$

where G_A and G_B are independent Gaussian matrices with nonsingular covariance matrices after vectorization. This is the usual asymptotic normality condition for correctly specified fixed-dimensional regression with nondegenerate coefficient estimation error.

For each fold k, let $\widehat{\Gamma}_{n,k}$ be the empirical covariance of $\phi(X)$ on $\mathcal{I}_\beta \setminus I_k$. Also assume that $\widehat{B}^A - \widehat{B}^B$ has row rank d_C and $\widehat{\Gamma}_{n,k} \succ 0$ almost surely for all sufficiently large sample sizes. Coefficient estimates with a joint density and calibration feature vectors in affine general position give these ranks when there are more than p calibration observations.

For every fixed $\tau > 0$, the disagreement covariance is

$$V_{n,k} = \frac{\tau^2}{4} (\widehat{B}^A - \widehat{B}^B) \widehat{\Gamma}_{n,k} (\widehat{B}^A - \widehat{B}^B)^\top.$$

The sampling assumptions give $\widehat{\Gamma}_{n,k} \to_P \Gamma$, so

$$n_C V_{n,k} \Rightarrow \frac{\tau^2}{4} (G_A - G_B) \Gamma (G_A - G_B)^\top.$$

The limiting matrix is positive definite almost surely: $p \geq d_C$ and the Gaussian coefficient difference has a density on $\mathbb{R}^{d_C \times p}$. Hence $\lambda_{\min}(V_{n,k})^{-1} = O_P(n_C)$. The coefficient expansion and finite feature second moments also give $\rho_n^2 = O_P(n_C^{-1})$. Since $V_{n,k}$ is full rank, the range condition holds, and conditional projection yields

$$\kappa_{n,k} \leq \frac{\mathbb{E}[\|\zeta_{n,k}\|^2 \mid \mathcal{T}_{n,k}]}{\lambda_{\min}(V_{n,k})} \leq \frac{\rho_n^2}{\lambda_{\min}(V_{n,k})} = O_P(1).$$

Thus Assumption 3(b) holds. Moreover, $\operatorname{tr}(V_{n,k}) = O_P(n_C^{-1})$, so part (a) holds as well.

C.2 Gaussian overlap and residual variation

Lemma 1 (Preservation of residual variation). *Fix a fold and condition on $\mathcal{T}_{n,k}$, suppressing k in the notation. Under Assumptions 2 and 3(b), let $Q_n = \mathbb{E}[\operatorname{Var}(X \mid U_n, E, \mathcal{T}_{n,k}) \mid \mathcal{T}_{n,k}]$. Draw X, X' independently conditional on the same (C, E), and put*

$$\Delta X = X - X', \qquad d_n^2 = \{\widehat{C}_n(X) - \widehat{C}_n(X')\}^\top V_n^\dagger \{\widehat{C}_n(X) - \widehat{C}_n(X')\},$$
$$B_n = \tfrac{1}{2}\mathbb{E}[\Delta X \Delta X^\top e^{-d_n^2/4} \mid \mathcal{T}_{n,k}].$$

Then $Q_n \succeq B_n$ and $\lambda_{\min}(B_n)^{-1} = O_P(1)$. For the average $\widetilde{m}_{n,X}$ of the conditional expectations $m_{n,X}$ over the M shared perturbations, the residual Gram matrix $G_{n,M} = \mathbb{E}[(X - \widetilde{m}_{n,X})(X - \widetilde{m}_{n,X})^\top \mid \mathcal{T}_{n,k}]$ satisfies

$$\lambda_{\min}(G_{n,M}) \geq \frac{\lambda_{\min}(B_n)^2}{\lambda_{\max}(\mathbb{E}[XX^\top])}, \qquad \|G_{n,M}^{-1}\| = O_P(1).$$

Proof. Fix $(C, E) = (c, e)$. By the range condition, the conditional Gaussian laws $U_n \mid X = x$ have the same affine support for almost every x in this conditional distribution. Let $\pi_x(u)$ be their densities on that support, and let $\pi(u) = \mathbb{E}[\pi_X(u) \mid c, e, \mathcal{T}_{n,k}]$ be the mixture density. The conditional-variance identity gives, for any $v \in \mathbb{R}^d$,

$$\mathbb{E}[\operatorname{Var}(v^\top X \mid U_n, c, e, \mathcal{T}_{n,k}) \mid c, e, \mathcal{T}_{n,k}] = \tfrac{1}{2}\mathbb{E}\left[(v^\top \Delta X)^2 \int \frac{\pi_X(u)\pi_{X'}(u)}{\pi(u)}\, du \,\middle|\, c, e, \mathcal{T}_{n,k}\right].$$

Cauchy–Schwarz and the affinity of Gaussians with common covariance imply

$$\int \frac{\pi_x \pi_{x'}}{\pi}\, du \geq \left(\int \sqrt{\pi_x \pi_{x'}}\, du\right)^2 = e^{-d_n^2/4}.$$

Conditioning additionally on C can only decrease conditional variance; integration over (C, E) gives $Q_n \succeq B_n$. The zero-rank case is immediate because the range condition then gives $\zeta_n = 0$.

Write $s = \lambda_{\min}(\Sigma_C) > 0$. The conditional centering of ζ_n gives $\mathbb{E}[d_n^2 \mid \mathcal{T}_{n,k}] = 2\kappa_n$, while $\frac{1}{2}\mathbb{E}[\Delta X \Delta X^\top] = \Sigma_C$. Given any $\epsilon > 0$, choose a fixed $A < \infty$ such that $\kappa_n \leq A$ with probability at least $1-\epsilon$ for all sufficiently large n. Since ΔX has a fixed distribution with finite second moment, choose $L < \infty$ with $\mathbb{E}[\|\Delta X\|^2 \mathbf{1}\{\|\Delta X\|^2 > L\}] < s/2$. On $\{\kappa_n \leq A\}$, Markov's inequality gives

$$\mathbb{E}[\|\Delta X\|^2 \mathbf{1}\{d_n^2 > T\} \mid \mathcal{T}_{n,k}] \leq s/2 + 2LA/T \leq s$$

for a sufficiently large fixed T. Hence $B_n \succeq e^{-T/4}(\Sigma_C - sI/2) \succeq e^{-T/4} sI/2$ on this event, proving the inverse bound.

For exact prediction averaging define $S_n X = \mathbb{E}_\xi[m_{n,X}(\widehat{C}_n(X) + \xi, E) \mid X, E, \mathcal{T}_{n,k}]$ and $G_{n,\infty} = \mathbb{E}[(X - S_n X)(X - S_n X)^\top \mid \mathcal{T}_{n,k}]$. Conditional-expectation identities yield $\mathbb{E}[X(X - S_n X)^\top \mid \mathcal{T}_{n,k}] = Q_n$. For any v, Cauchy–Schwarz therefore gives

$$\|Q_n v\|^2 \leq \lambda_{\max}(\mathbb{E}[XX^\top])\, v^\top G_{n,\infty} v, \qquad \lambda_{\min}(G_{n,\infty}) \geq \frac{\lambda_{\min}(Q_n)^2}{\lambda_{\max}(\mathbb{E}[XX^\top])}.$$

Finally, $\mathbb{E}[\widetilde{m}_{n,X} \mid X, E, \mathcal{T}_{n,k}] = S_n X$, so $G_{n,M} = G_{n,\infty} + \mathbb{E}[\operatorname{Var}(\widetilde{m}_{n,X} \mid X, E, \mathcal{T}_{n,k}) \mid \mathcal{T}_{n,k}] \succeq G_{n,\infty}$. □

D PROOFS

D.1 PROOF OF THEOREM 1

Proof. Recall from § 2.3 that $Y = \beta_0^\top X + q_E(X) + W$ with $\mathbb{E}[W \mid X, E] = 0$. Under IR, $q_E(X)$ is measurable with respect to $(C(X), E)$, and $\mathbb{E}[W \mid C(X), E] = 0$ by the tower property. Therefore, $\widetilde{Y} = \beta_0^\top \widetilde{X} + W$, and it follows that $\mathbb{E}[\widetilde{X}\widetilde{Y}] = \Sigma_C \beta_0$. Positivity completes the proof. □

D.2 BIAS WITHOUT THE INCLUSION RESTRICTION

Under the setting of § 2.1, define

$$r_E(X) \coloneqq q_E(X) - \mathbb{E}[q_E(X) \mid C(X), E].$$

Subtracting the conditional expectation given $(C(X), E)$ from $Y = \beta_0^\top X + q_E(X) + W$ gives

$$\widetilde{Y} = \beta_0^\top \widetilde{X} + r_E(X) + W, \qquad \mathbb{E}[\widetilde{X}\widetilde{Y}] = \Sigma_C \beta_0 + \mathbb{E}[\widetilde{X} r_E(X)],$$

where $\mathbb{E}[\widetilde{X} W] = 0$ by the tower property. Under positivity, multiplying by Σ_C^{-1} yields

$$\beta_{\mathrm{OVR}} - \beta_0 = \Sigma_C^{-1} \mathbb{E}[\widetilde{X} r_E(X)]. \tag{13}$$

For the sensitivity bound in § 3.1, suppose $q_{e_0}(X) \in L^2$. Since $q_E(X) - q_{e_0}(X)$ is determined by $(C(X), E)$,

$$r_E(X) = q_{e_0}(X) - \mathbb{E}[q_{e_0}(X) \mid C(X), E].$$

The least-squares property of conditional expectation gives

$$\mathbb{E}[r_E(X)^2] \le \mathbb{E}\Big[\big(q_{e_0}(X) - \mathbb{E}[q_{e_0}(X) \mid C(X)]\big)^2\Big] = \delta_{\mathrm{IR}}^2.$$

For every unit vector $v \in \mathbb{R}^d$, (13) and Cauchy–Schwarz imply

$$\big|v^\top(\beta_{\mathrm{OVR}} - \beta_0)\big| \le \{v^\top \Sigma_C^{-1} v\}^{1/2} \{\mathbb{E}[r_E(X)^2]\}^{1/2}.$$

Taking the supremum over unit vectors yields

$$\|\beta_{\mathrm{OVR}} - \beta_0\|_2 \le \frac{\{\mathbb{E}[r_E(X)^2]\}^{1/2}}{\sqrt{\lambda_{\min}(\Sigma_C)}} \le \frac{\delta_{\mathrm{IR}}}{\sqrt{\lambda_{\min}(\Sigma_C)}}.$$

Equality without IR. Let H, U_X, U_Y be mutually independent standard normal variables, independent of E. For constants $a_e \in \mathbb{R}$, consider the SCM

$$X = H + U_X, \qquad Y = \beta_0 X + H^2 - 1 + a_E + U_Y.$$

Since $\mathbb{E}[H^2 \mid X = x] = 1/2 + x^2/4$, we have

$$q_e(x) = a_e + \frac{x^2 - 2}{4}, \qquad C_e(x) = a_e - a_{e_0}.$$

Thus $C(X)$ is constant, $\widetilde{X} = X$, and $\Sigma_C = 2 > 0$. The remaining nuisance is

$$r_E(X) = \frac{X^2 - 2}{4},$$

which is nonzero almost surely, so IR fails. Nevertheless, symmetry of X gives

$$\mathbb{E}[\widetilde{X} r_E(X)] = \frac{1}{4}\big(\mathbb{E}[X^3] - 2\mathbb{E}[X]\big) = 0.$$

The bias identity therefore yields $\beta_{\mathrm{OVR}} = \beta_0$ despite the violation of IR: the remaining nuisance is nonconstant but uncorrelated with the residualized treatment.

D.3 PROOF OF PROPOSITION 2

Proof. If $e_\star = e_0$, then $q_{e_0}(X)$ is constant. Otherwise, $C_{e_\star}(X) = \kappa_\star - q_{e_0}(X)$, so $q_{e_0}(X)$ is determined by $C(X)$. Thus IR holds, and Theorem 1 gives the result. □

D.4 Proof of Proposition 3

Proof. Conditional independence and recoverability give $q_e(X) = \rho_e(Z_C)$, and hence $C(X) = \Delta(Z_C)$. Injectivity makes Z_C measurable with respect to $C(X)$, proving IR, and also gives $\Sigma_C = \mathbb{E}[\mathrm{Var}(X \mid Z_C, E)] = \Sigma_{Z_C}$. Positivity and Theorem 1 complete the proof. □

D.5 Proof of Theorem 3

Proof. Fix a fold, condition on $\mathcal{T}_{n,k}$, and suppress k. All expectations in the first three steps are over a fresh observation and the prediction perturbations, conditional on this training data.

Step 1: recovery of the confounding term. Under IR, $q_E(X) = \varphi_E(C(X))$ is square integrable. For every $\epsilon > 0$, choose bounded Lipschitz functions b_e such that $\mathbb{E}[(q_E(X) - b_E(C(X)))^2] < \epsilon$; such functions are dense in L^2 of the finite-dimensional contrast distribution, separately in each of the finitely many environments. Let L_b bound their Lipschitz constants. Because $\mathbb{E}\|U_n - C(X)\|^2 = \rho_n^2 + \mathrm{tr}(V_n)$, the projection property gives

$$\mathbb{E}[(q_E - m_{n,q}(U_n, E))^2] \le \mathbb{E}[(q_E - b_E(U_n))^2] \le 2\epsilon + 2L_b^2\{\rho_n^2 + \mathrm{tr}(V_n)\}.$$

Part (a) of Assumption 3, followed by $\epsilon \downarrow 0$, makes this $o_P(1)$. For the average $\widetilde{m}_{n,q} = M^{-1}\sum_j m_{n,q}(\widehat{C}_n(X) + \xi_j, E)$, Jensen's inequality and the common Gaussian marginals give

$$\mathbb{E}[(q_E - \widetilde{m}_{n,q})^2] \le \frac{1}{M}\sum_j \mathbb{E}[(q_E - m_{n,q}(\widehat{C}_n(X) + \xi_j, E))^2] = o_P(1). \tag{14}$$

Step 2: the population residual equation. Linearity of conditional expectation and $\mathbb{E}[W \mid X, E] = 0$ imply $m_{n,Y} = \beta_0^\top m_{n,X} + m_{n,q}$. Using the same prediction perturbations for X and Y preserves this identity after averaging. Thus the ideal residuals satisfy

$$R_X = X - \widetilde{m}_{n,X}, \qquad R_Y = Y - \widetilde{m}_{n,Y} = \beta_0^\top R_X + (q_E - \widetilde{m}_{n,q}) + W.$$

Since the perturbations are independent of W conditional on (X, E), $\mathbb{E}[R_X W] = 0$. By Lemma 1, $G_{n,M} = \mathbb{E}[R_X R_X^\top]$ has inverse norm $O_P(1)$. Moreover, $\mathbb{E}\|R_X\|^2 \le 4\mathbb{E}\|X\|^2$, by the L^2 contraction property and Jensen's inequality. Therefore the score $\mathbb{E}[R_X(R_Y - \beta_0^\top R_X)]$ vanishes by Eq. (14) and Cauchy–Schwarz. The population slope $\beta_{n,M} = G_{n,M}^{-1}\mathbb{E}[R_X R_Y]$ obeys the sharper bound

$$\|\beta_{n,M} - \beta_0\| \le \frac{\|q_E - \widetilde{m}_{n,q}\|_{L^2}}{\sqrt{\lambda_{\min}(G_{n,M})}} = o_P(1).$$

If the φ_e are globally L-Lipschitz, the numerator is at most $L\sqrt{\rho_n^2 + \mathrm{tr}(V_n)}$, by comparing the noisy regression with $\varphi_E(U_n)$ in Step 1.

Step 3: empirical ideal moments. Conditional expectations of a fixed square-integrable target have uniformly integrable squared norms, uniformly over the conditioning variables: truncate the target at a fixed level, use the L^2 contraction bound for its tail, then let that level grow. This applies conditionally on the independent training data, since the fresh target distribution is fixed. Finite averages of predictions inherit the bound, as do the squared ideal residuals and their products with W and $q_E - \widetilde{m}_{n,q}$. Conditional independence across held-out rows and truncation therefore give foldwise weak laws for the ideal Gram and score moments. With environment-stratified folds, apply the same argument within each environment and weight by the limiting environment proportions. Pooling the fixed number of folds gives

$$\frac{1}{n}\sum_i R_{X,i}R_{X,i}^\top - \sum_k \frac{|I_k|}{n} G_{n,k,M} = o_P(1), \qquad \frac{1}{n}\sum_i R_{X,i}(R_{Y,i} - \beta_0^\top R_{X,i}) = o_P(1).$$

By Lemma 1 and fixed K, the convex combination of Gram matrices has inverse norm $O_P(1)$.

Step 4: transfer to fitted regressions. Let $\nu_{Z,i}$ be the average of $f_{n,k,Z} - m_{n,k,Z}$ over the prediction draws for row $i \in I_k$. Part (c) of Assumption 3 and Jensen's inequality give $\mathbb{E}[\|\nu_{Z,i}\|^2 \mid \mathcal{T}_{n,k}] = o_P(1)$. Conditional Markov's inequality then implies $n^{-1}\sum_i \|\nu_{Z,i}\|^2 = o_P(1)$ for $Z = X, Y$.

Since $\widehat{X}_i = R_{X,i} - \nu_{X,i}$ and $\widehat{Y}_i = R_{Y,i} - \nu_{Y,i}$, empirical Cauchy–Schwarz shows that replacing the ideal residuals by the fitted ones changes the Gram and score moments by $o_P(1)$. Consequently,

$$\Pr(\widehat{G}_n \succ 0) \to 1, \quad \|\widehat{G}_n^{-1}\| = O_P(1), \quad \widehat{s}_n := n^{-1} \sum_i \widehat{X}_i(\widehat{Y}_i - \beta_0^\top \widehat{X}_i) = o_P(1).$$

Finally, $\widehat{\beta}_{\mathrm{SOVR}} - \beta_0 = (\widehat{G}_n + \lambda_n I)^{-1}(\widehat{s}_n - \lambda_n \beta_0) = o_P(1)$. □

D.6 Proof of Proposition 1

We first prove a generalization of Proposition 1 that allows environment-specific intercepts and does not require a specified latent variable.

Proposition 4 (General low-rank inclusion restriction)**.** *Under the setting of § 2.1, let $g : \mathbb{R}^d \to \mathbb{R}^r$ be a measurable function. Suppose that for every $e \in \mathcal{E}$ there are an intercept $\alpha_e \in \mathbb{R}$ and loadings $\lambda_e \in \mathbb{R}^r$ such that*

$$q_e(X) = \alpha_e + \lambda_e^\top g(X) \qquad P_X\text{-almost surely}.$$

Fix a reference environment e_0, define $L_{e,\cdot} := (\lambda_e - \lambda_{e_0})^\top$ for $e \neq e_0$, and set $\Sigma_g := \mathbb{E}[\mathrm{Var}(X \mid g(X), E)]$. If $\lambda_{e_0} \in \mathrm{rowspan}(L)$ and $\Sigma_g \succ 0$, then Assumptions 1 and 2 hold and $\beta_{\mathrm{OVR}} = \beta_0$.

Proof of Proposition 4.

Let $\delta_\alpha := \big(\alpha_e - \alpha_{e_0}\big)_{e \neq e_0}$. The contrast satisfies $C(X) = \delta_\alpha + Lg(X)$. Since $\lambda_{e_0} \in \mathrm{rowspan}(L)$, there exists a $b \in \mathbb{R}^{|\mathcal{E}|-1}$ such that $\lambda_{e_0} = L^\top b$. It follows that

$$q_{e_0}(X) = \alpha_{e_0} + b^\top \big(C(X) - \delta_\alpha\big),$$

which proves IR. Moreover, since $C(X)$ is a function of $g(X)$, the conditional variance decomposition gives

$$\Sigma_C = \mathbb{E}[\mathrm{Var}(X \mid C(X), E)] \succeq \mathbb{E}[\mathrm{Var}(X \mid g(X), E)] = \Sigma_g \succ 0.$$

Thus positivity holds and Theorem 1 gives $\beta_{\mathrm{OVR}} = \beta_0$. □

Proof of Proposition 1. In (11), define $g(x) := \mathbb{E}[H \mid X = x]$ and set $\alpha_e = 0$ for every e. The common conditional expectation $\mathbb{E}[H \mid X, E] = \mathbb{E}[H \mid X]$ and $\mathbb{E}[U_Y \mid X, E] = 0$ imply $q_e(X) = \lambda_e^\top g(X)$. The two conditions in Proposition 1 are therefore those of Proposition 4, which proves the claim. □

E Variance correction for OVR (VOVR)

When positivity fails, OVR does not identify β_0 in the directions of X determined by $C(X)$. Under additional restrictions on the conditional outcome variances, these directions can be recovered from cross-environment variance contrasts.

Population identification. Suppose the conditional nuisances have the low-rank form of Proposition 4 and that

$$v_e(x) := \mathrm{Var}(Y \mid X = x, E = e) = \sigma^2(x) + \lambda_e^\top \Sigma_H \lambda_e,$$

where $\Sigma_H \succ 0$ is constant in (x, e) and $\sigma^2(x)$ is environment invariant. In the latent outcome model, this holds if $\mathbb{E}[H \mid X, E] = \mathbb{E}[H \mid X]$, $\mathrm{Var}(H \mid X, E) = \Sigma_H$ and the outcome noise is conditionally uncorrelated with variance $\sigma^2(x)$. Assume the loading differences span all r latent directions and the centered contrasts have rank r. Choose r independent centered reference contrasts $z(x)$; then

$$C_e(x) - \mathbb{E}[C_e(X)] = \omega_e^\top z(x), \qquad q_{e_0}(x) = \alpha_0' + b^\top z(x), \qquad \omega_{e_0} = 0.$$

Let $\Sigma_z \succ 0$ denote the latent covariance expressed in this basis. The variance contrasts cancel the common noise and satisfy

$$C_e^v := v_e(x) - v_{e_0}(x) = \omega_e^\top \Sigma_z \omega_e + 2\omega_e^\top u, \qquad u = \Sigma_z b. \tag{15}$$

These are linear equations in $\mathrm{vech}(\Sigma_z)$ and u, with doubled off-diagonal quadratic coefficients. If the stacked design has full column rank, which requires $|\mathcal{E}| - 1 \geq r(r+3)/2$, the system determines

Table 2: Variance correction on Gaussian controls ($n = 20{,}000$, ten replications). Entries are mean ℓ_2 estimation error (Monte Carlo s.e.) [full estimates/attempts], averaged over accepted fits; — denotes no accepted fit.

Setting	SOVR	Hard hybrid	Plain VOVR
Affine, $r = 1$	1.709(0.077) [10/10]	0.656(0.141) [10/10]	0.476(0.058) [10/10]
Affine, $r = 2$	1.343(0.084) [10/10]	0.484(0.063) [8/10]	0.493(0.066) [8/10]
(a-G), nonlinear	0.199(0.027) [10/10]	0.195(0.029) [8/10]	— [0/10]
(a-G), affine	1.076(0.069) [10/10]	0.614(0.161) [4/10]	0.621(0.157) [4/10]

(Σ_z, u) and $b = \Sigma_z^{-1}u$. Subtracting $C_E(X) + b^\top z(X)$ from Y leaves $\beta_0^\top X$ plus an environment intercept and a conditionally mean-zero error. Consequently, OLS on X and environment indicators identifies β_0 whenever $\Sigma_E \coloneqq \mathbb{E}[\mathrm{Var}(X \mid E)] \succ 0$, even if Σ_C is singular, for example for affine $g(X) = AX$.

Feasible hybrid. Plain VOVR estimates the rank r, the loading coordinates ω_e and the variance functions v_e on $\mathcal{I}_C$, solves a weighted ridge version of (15) for $(\widehat{\Sigma}_z, \widehat{u})$, and regresses $Y_i - \widehat{C}_{E_i}(X_i) - \widehat{b}^\top \widehat{z}(X_i)$ on X_i and environment indicators on $\mathcal{I}_\beta$ to obtain $\widehat{\beta}_{\mathrm{VOVR}}$. The fit is rejected unless the variance contrasts are jointly nonzero and constant in x up to estimation noise, their system is full rank and well conditioned, and $\lambda_{\min}(\widehat{\Sigma}_z)$ exceeds its estimated noise floor. The hard hybrid applies this correction only where OVR is weak. It compares each eigenvalue of the whitened SOVR residual covariance with a noise floor estimated from the disagreement of the two contrast fits, collects the directions below the floor in a projector $\Pi^\perp$, and returns

$$\widehat{\beta}_{\mathrm{hard}} = (I - \Pi^\perp)\widehat{\beta}_{\mathrm{OVR}} + \Pi^\perp \widehat{\beta}_{\mathrm{VOVR}};$$

if no weak direction is detected, it returns $\widehat{\beta}_{\mathrm{OVR}}$, and if the variance fit is rejected, it returns only the identified component $(I - \Pi^\perp)\widehat{\beta}_{\mathrm{OVR}}$.

Empirical comparison. We evaluate both variants on four Gaussian controls with ten replications at $n = 20{,}000$, sharing the SOVR contrasts, splits and slope folds. The affine controls use $X \sim N(0, I_{r+2})$, $H \mid X \sim N(AX, \Sigma_H)$ with $A = [I_r, 0]$, and $Y = \beta_0^\top X + \alpha_E + \lambda_E^\top H + .5U_Y$, with three environments for $r = 1$ and six for $r = 2$, so positivity fails in r directions. The (a-G) controls modify panel (a) to have Gaussian latents, $H \mid X \sim N(g(X), \Sigma_H)$ with six environments, using either its nonlinear posterior mean g or an affine replacement. Table 2 reports the results.

F Experimental Details

F.1 Synthetic SCMs

Environments are uniform, and all exogenous noises (U_X, U_H, U_R, U_Y) are mutually independent standard Gaussians of the appropriate dimension, independent of E; the outcome noise enters as $.5U_Y$. The latent-loading and recoverable-state experiments hold one structural law fixed while redrawing observations. The MLP experiments average over ten independently drawn structural laws, with the law for each replication held fixed across sample sizes. Master seed 20260807 generates panels (a)–(c), and seed 20260808 generates the paired panels (d)–(f). Every random MLP has width 32 and tanh activations: for hidden layer ℓ, $a_\ell = \tanh(a_{\ell-1}W_\ell + b_\ell)$, with entries of W_ℓ drawn from $N(0, 1/d_{\ell-1})$ and entries of b_ℓ from $N(0, .1^2)$; the output layer is linear with the same initialization scaling.

Latent loadings. Let $H_1, H_2 \overset{\mathrm{iid}}{\sim} \mathrm{Bernoulli}(1/2)$, $d = 4$, and split X into two blocks $X^{(j)} \in \mathbb{R}^2$. The unit vectors a_j are normalized Gaussian draws, and $X^{(j)} \mid H_j = h \sim N((2h-1).8a_j, \sigma_h^2 I_2)$, where $\sigma_0 = .7$ and $\sigma_1 = 1.3$. We draw $E \sim \mathrm{Unif}\{0, 1, 2, 3\}$ independently of (H, X) and set $Y = \beta_0^\top X + \lambda_E^\top H + .5U_Y$, with $\beta_0 \sim N(0, I_4)$. Gaussian QR draws determine the left and right singular vectors of the loading contrasts; before multiplication by the confounding scale 2, their singular values are 1.5 and 1, and the reference loading has norm 1. Thus the realized 3×2 contrast-loading matrix has singular values 3 and 2, and the realized reference loading has norm 2. Hence

$C(X) = L\mathbb{E}[H \mid X]$ obeys the rank condition exactly; unequal block variances make the posterior nonlinear, while two-dimensional blocks leave residual treatment variation.

One unconfounded environment. Here $d = 4$, $d_H = 2$, and $E \sim \mathrm{Unif}\{0, \dots, 4\}$. Draw $m_e = .3z_e$ and $b_e = .5v_e$ with $z_e \sim N(0, I_2)$ and $v_e \sim N(0, I_4)$, then set $H = m_E + U_H$ and $X = \mathrm{std}\{f_X(H)\} + b_E + U_X$, where $f_X : \mathbb{R}^2 \to \mathbb{R}^4$ is a two-hidden-layer MLP and std denotes coordinatewise centering and scaling on an independent calibration sample of size 4,000. A second two-hidden-layer MLP maps the concatenated pair $(\mathrm{onehot}(E), H)$ to a scalar, which is pooled-standardized to $\widetilde{\psi}(H, E)$. We use $Y = \beta_0^\top X + 3\widetilde{\psi}(H, E)\mathbf{1}\{E \neq 3\} + .5U_Y$, with $\beta_0 \sim N(0, I_4)$. Thus environment 3 is exactly clean, while the Gaussian U_X gives all environments common support.

Recoverable confounding state. Let $E \sim \mathrm{Unif}\{0, 1, 2, 3\}$, $R = \exp(.6U_R - .18)$, and $X = RV_E \in \mathbb{R}^4$ with $\|V_e\| = 1$. Relative to the uniform distribution on the unit sphere, V_e has density $1 + .5\theta_e^\top v$, where $\theta_e = (\cos(2\pi e/4), \sin(2\pi e/4), 0, 0)$. Set $b_e = 3\cos(2\pi e/4)$, $(d_0, d_1, d_2, d_3) = (-.5, -1/6, 1/6, .5)$, $H = R^2 + b_E R + .5U_H$, and $Y = \beta_0^\top X + H + d_E + .5U_Y$, with $\beta_0 \sim N(0, I_4)$. This is the example following Proposition 3 with latent state $Z_C = R = \|X\|$, and $X \perp\!\!\!\perp H \mid (R, E)$. Then $q_e(x) = \|x\|^2 + b_e\|x\| + d_e$, the contrast is injective in R, and the exact propensity lower bound is $(1 - .5)/4 = .125$. Angular variation supplies positivity.

Shared decoder: outcome, treatment, and latent shifts. Panels (d)–(f) use $d = 4$, $d_H = 1$, and five environments. Let $\widetilde{A}_e \in \mathbb{R}^{d_H}$ and $\widetilde{B}_e \in \mathbb{R}^d$ be independent standard Gaussian draws, centered across environments and divided by the standard deviation of all their entries. With independent three-hidden-layer MLPs f_H, f_X, and $\psi_e : \mathbb{R} \to \mathbb{R}$, we generate $H = \mathrm{std}\{f_H(U_H)\} + t\widetilde{A}_E$, $X = \mathrm{std}\{f_X(H, U_X)\} + s\widetilde{B}_E$, and $Y = \beta_0^\top X + 3\,\mathrm{std}\{\psi_E(H)\} + .5U_Y$, where $\|\beta_0\|_2 = 2$ and standardization uses 10,000 independent calibration draws. The shift pairs (t, s) are $(0, 0)$, $(0, .3)$, and $(.3, .3)$, yielding Outcome shifts, Outcome and treatment shifts, and Latent, outcome and treatment shifts. Within each replication the three settings share structural maps, exogenous draws, and learner seeds. The first has $E \perp\!\!\!\perp (H, X)$; the latter two successively add $E \to X$ and $E \to H$ while retaining common latent and treatment mechanisms.

F.2 Estimators and synthetic results

Both OVR estimators use an environment-balanced 50–50 contrast/slope split and three-fold cross-fitting on the slope split. The first- and second-stage networks have two hidden layers of width 64 and train for at most 200 and 300 epochs. DOVR fits one shared-trunk outcome network on the full contrast split, uses one head per non-reference environment with fixed penalty 10^{-4}, and passes the resulting full contrast directly to the second stage. SOVR divides the contrast split again, fits two members with penalty 10^{-2} and 15% validation sets, averages their contrasts, and uses their disagreement covariance for Gaussian smoothing with $\tau = 1$ and $M = 8$ perturbations arranged in four antithetic pairs.

OLS ignores E. FE removes environment means from X and Y before regression. Anchor regression uses environment indicators and $\gamma = 2$. In designs with $E \perp\!\!\!\perp X$, the population anchor transform supplies no covariate signal, explaining why anchor, OLS, and FE nearly coincide. All plotted points are mean ℓ_2 estimation error; shaded bands in Fig. 5 show $\pm$ one Monte Carlo standard error.

Linear-plus-neural-network baseline (LNN). LNN jointly optimizes (β, θ) to minimize squared prediction error for $\beta^\top X + f_\theta(X, E)$. The neural network f_θ receives standardized X and environment indicators and has two width-64 ReLU layers. The linear coefficient is unpenalized. Validation outcome error selects network weight decay from $\{0, 10^{-4}, 10^{-2}\}$, initialization of β from FE or zero, and the stopping epoch, with patience 30. Adam uses learning rate .001 and batches of 1,024; training lasts at most 200 epochs on synthetic data and 120 on real data.

For the reported comparison, 20% of observations are reserved for a prediction check. An inner validation split of the remaining observations selects the settings, after which the model is refitted on the full sample for coefficient evaluation. Splits keep each user in one subset when user identifiers are available. Standardization is fitted on the corresponding training sample. Since $f_\theta(X, E)$ can represent linear dependence on X, the prediction objective alone does not identify β; LNN reports the coefficient selected by this training procedure.

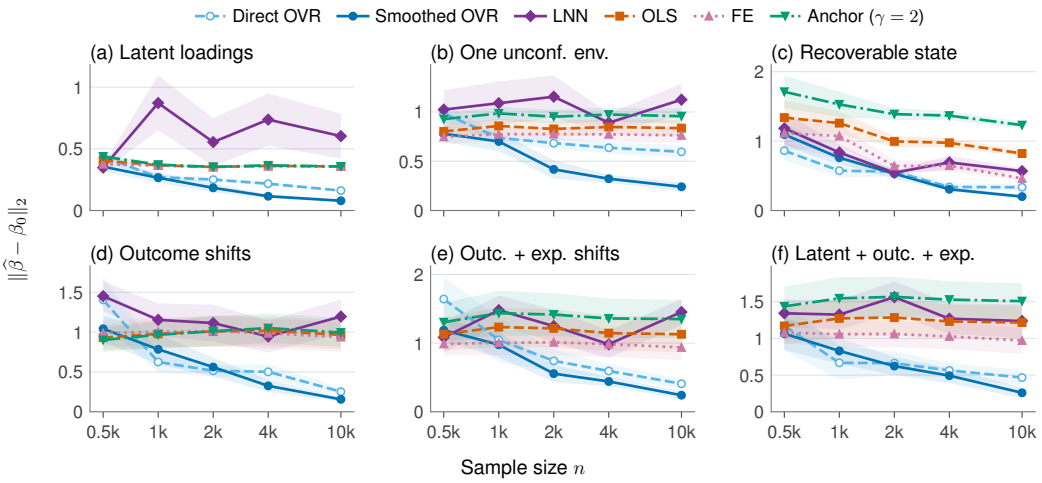


Figure 6: Synthetic comparison at $nn \in \{0.5, 1, 2, 4, 10\}$k, including Direct OVR. Means and $\pm$ one Monte Carlo s.e. over ten paired replications.

F.3 Supplementary synthetic comparisons

Fig. 6 adds DOVR to the main comparison at $n \in \{500, 1000, 2000, 4000, 10000\}$, using the same observations and ten replications as Fig. 5.

F.4 Controlled component ablation

We compare five nested variants at $n = 10{,}000$ in all six synthetic settings, with ten paired replications per setting: DOVR with one contrast fit, the same fit with stronger contrast-head regularization, the regularized fit with a 15% validation holdout and early stopping, two such fits on disjoint halves of the contrast split averaged without smoothing ($\tau = 0$), and the same ensemble with uncertainty smoothing ($\tau = 1$). All variants share observations, the outer split, second-stage folds, and nuisance-learner seeds. The last two variants also share the fitted ensemble and disagreement estimate, so their comparison isolates smoothing. For pairing, the ablation DOVR uses SOVR's second-stage folds and can therefore differ from DOVR run separately. The SOVR estimates match those in Fig. 5 at $n = 10{,}000$.

Let $n_\beta = |\mathcal{I}_\beta|$ be the number of observations used to estimate the coefficient, and define the residual second moment and cross-moment

$$\widehat{G} = \frac{1}{n_\beta} \sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{X}_i^\top, \qquad \widehat{h} = \frac{1}{n_\beta} \sum_{i \in \mathcal{I}_\beta} \widehat{X}_i \widehat{Y}_i.$$

The implementation uses $\widehat{G}_\lambda = \widehat{G} + \lambda I$ with a numerical ridge $\lambda = 10^{-10}$, so

$$\widehat{\beta} - \beta_0 = \widehat{G}_\lambda^{-1} \widehat{r}, \qquad \widehat{r} := \widehat{h} - \widehat{G}_\lambda \beta_0.$$

Thus $\|\widehat{r}\|_2$ measures the violation of the empirical estimating equation at β_0, while $\|\widehat{G}_\lambda^{-1}\widehat{r}\|_2 / \|\widehat{r}\|_2$ measures how much the inverse residual second-moment matrix scales this violation. Their product is the ℓ_2 estimation error. The latter factor concerns the observed direction $\widehat{r}$, rather than the operator norm of $\widehat{G}_\lambda^{-1}$.

Table 3: Component ablation at $n = 10{,}000$. For each displayed quantity, entries are geometric means of the added-to-previous variant ratios over 60 paired comparisons (six settings, ten replications each); values below one indicate reductions. "Lower error" counts pairs in which the added component reduces $\|\widehat{\beta} - \beta_0\|_2$.

Added component	$\|\widehat{\beta} - \beta_0\|_2$	Lower error	$\|\widehat{r}\|_2$	$\frac{\|\widehat{G}_\lambda^{-1}\widehat{r}\|_2}{\|\widehat{r}\|_2}$
Contrast-head regularization	0.926	42/60	0.774	1.196
Validation and early stopping	0.656	42/60	0.611	1.074
Disjoint split ensemble ($\tau = 0$)	1.080	25/60	0.881	1.227
Uncertainty smoothing ($\tau = 1$)	0.792	44/60	1.147	0.691

Validation and early stopping give the largest aggregate reductions in estimation error and $\|\widehat{r}\|_2$. The disjoint ensemble alone increases geometric-mean estimation error by 8.0%, while providing the disagreement estimate used for smoothing. Relative to this identical unsmoothed ensemble, smoothing reduces geometric-mean estimation error by 20.8%: $\|\widehat{r}\|_2$ increases by 14.7%, but $\|\widehat{G}_\lambda^{-1}\widehat{r}\|_2/\|\widehat{r}\|_2$ decreases by 30.9%.

Smoothing increases $\lambda_{\min}(\widehat{G})$ in $58/60$ pairs and the positivity diagnostic $\widehat{\gamma}$ (§ B.1) in $59/60$. In each of the six settings, smoothing yields lower arithmetic-mean estimation error and a paired geometric-mean error ratio below one.

F.5 Sensitivity to the smoothing scale

We vary the smoothing scale over $\tau \in \{0, 0.25, 0.5, 1, 2, 4\}$ in all six synthetic settings at $n \in \{1{,}000, 10{,}000\}$, with ten paired replications per setting and sample size. Each replication fits the two contrast learners once and reuses their average, disagreement estimate, slope folds, and second-stage random seeds at every scale. All other settings match SOVR in § F.2. The $\tau = 0$ endpoint uses the same two-model ensemble without smoothing.

Fig. 7 summarizes the 720 fits. At $n = 1{,}000$, $\tau = 2$ has the lowest observed mean error in all six settings, although its paired differences from $\tau = 1$ are small relative to their Monte Carlo uncertainty. At $n = 10{,}000$, the lowest means occur at $\tau = 2$ in (a),(c), $\tau = 1$ in (b),(d), and $\tau = 0.5$ in (e),(f). The default $\tau = 1$ has lower mean error than $\tau = 0$ in all twelve setting–sample-size combinations. Larger perturbations can increase error: in (f) at $n = 10{,}000$, mean error rises from 0.259 at $\tau = 1$ to 0.537 at $\tau = 4$. The sweep supports $\tau = 1$ as a fixed default across settings.

F.6 KuaiRand-Pure held-out randomized reference

We restrict both assignment policies to a single recommendation setting (`tab`=1 in the dataset). The cohort, training/reference splits and reference regression are described in § 6.2. Both OVR estimators use a user-disjoint 50–50 contrast/slope split, width-64 two-layer networks trained for 120 fixed first-stage and 160 second-stage epochs, three second-stage folds, and batch size 512. DOVR fits one contrast network on the full contrast split with penalty 10^{-4}; SOVR fits two with penalty 10^{-2} and uses $\tau = 1$ with $M = 8$ perturbations arranged in four antithetic pairs. Across the ten training/reference splits, DOVR and SOVR have mean distances .0202 (.0022) and .0188 (.0013), versus .1289 (.0010) for OLS, .1039 (.0009) for environment fixed effects, and .1708 (.0010) for anchor regression; both OVR variants win all ten paired comparisons with each baseline. Their mean (minimum) positivity diagnostics $\widehat{\gamma}$ are .339 (.287) and .397 (.351), respectively. LNN uses the same training observations and scoring references, with user-disjoint validation; its mean reference distance is .095 (Table 1).

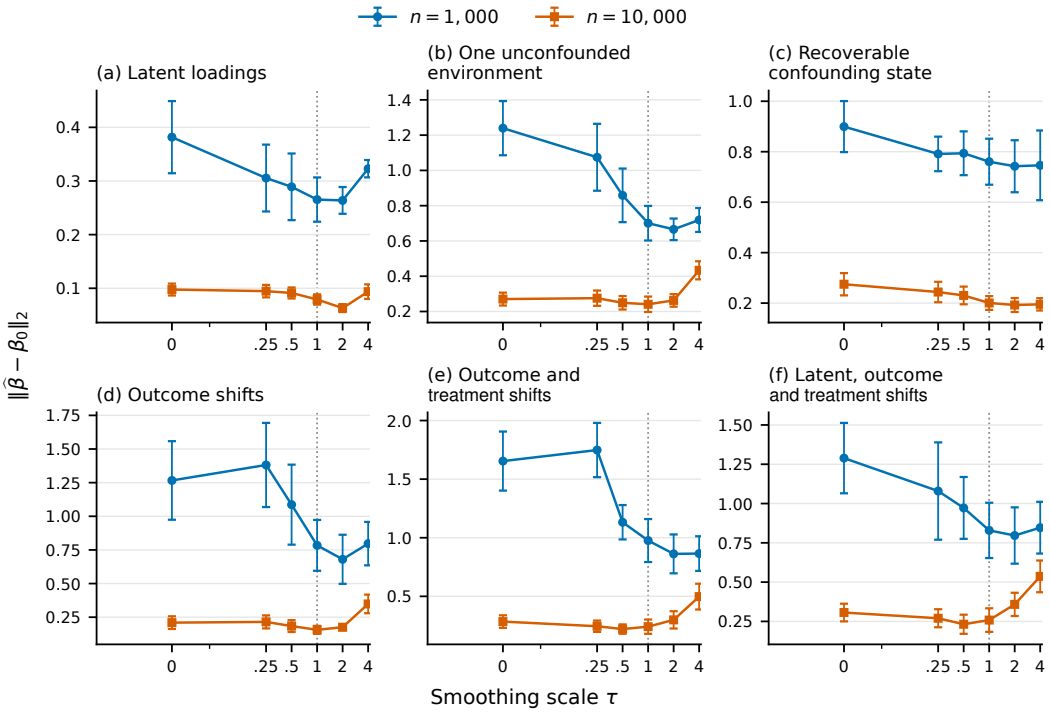


Figure 7: Sensitivity of Smoothed OVR to the smoothing scale. Points and bars show mean ℓ_2 estimation error and $\pm$ one Monte Carlo standard error over ten paired replications. All scales reuse the same fitted contrast ensemble, slope folds, and second-stage random seeds. The dotted vertical line marks the default $\tau = 1$; panels use separate linear error scales.