ESTIMATING CAUSAL EFFECT TRAJECTORIES UNDER DYNAMIC NETWORK INTERFERENCE

Anonymous authors
Paper under double-blind review

ABSTRACT

Dynamic network interference combines unknown propagation strengths and delays with treatment–outcome feedback. We propose **DyNICL** (Dynamic Network Interference Causal Learning) to estimate direct, peer, total, and interaction effect trajectories. Four fixed stochastic policies define the effects. Learned hop–lag exposure and history representations support neural backward-Q regression, with Q-based and assignment-weighted augmented estimation. Under sequential exchangeability, positivity, and exposure sufficiency, we identify the policy values and establish Q-based consistency under continuation convergence. For augmented estimation, we derive a cross-fitted orthogonal expansion separating continuation–assignment product error from conditional exposure drift. DyNICL-Q achieves lower peer and total effect RMSE than adapted NetEst and TNet across all nine synthetic neighbor-count/lag settings. On two semi-synthetic contact-network settings, DyNICL-Aug achieves the lowest peer and total effect RMSE, while DyNICL-Q achieves the lowest direct effect RMSE.

1 INTRODUCTION

In a networked system, a unit's outcome depends on both its own treatment and its peers' treatments (Hudgens & Halloran, 2008; Aronow & Samii, 2017). Repeated decisions extend these channels across time: earlier peer treatments affect current outcomes, which shape later assignments.

Two coupled problems obstruct estimation (Figure 1). Observed connections leave propagation strengths and delays unspecified, so fixed neighborhood averages merge treatments with different outcome contributions. Feedback creates a second difficulty: intervening on peer assignments changes both exposure and the histories governing subsequent treatments. Estimating a policy effect therefore requires following this evolving assignment process.

Longitudinal g-methods identify sequential policy effects (Robins, 1986; Hernán & Robins, 2010), augmented estimators combine regression with assignment weighting (Bang & Robins, 2005; Jiang & Li, 2016), and network representations encode interference (Forastiere et al., 2022; Guo et al., 2020; Wu et al., 2025). DyNICL connects these principles by learning hop–lag exposure and focal/peer history representations for neural backward-Q, supporting Q-based and augmented estimation of longitudinal policy effects.

Our contributions are:

- We define direct, peer, total, and interaction effect trajectories through attainable partial assignment regimes and identify their values by backward recursion, integrating non-intervened assignments under their natural law.
- We turn this recursion into a shared learned-exposure framework with Q-based and augmented estimators. Under representation sufficiency and nuisance convergence, we establish Q-based consistency and an augmented orthogonal expansion separating continuation–assignment product error, conditional exposure drift, and numerical integration error.
- DyNICL-Q lowers peer and total RMSE relative to NetEst and TNet across nine synthetic settings; DyNICL-Aug leads on these effects in both semi-synthetic cases. Ablations assess module contributions; perturbations assess graph and assignment-support sensitivity.

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.

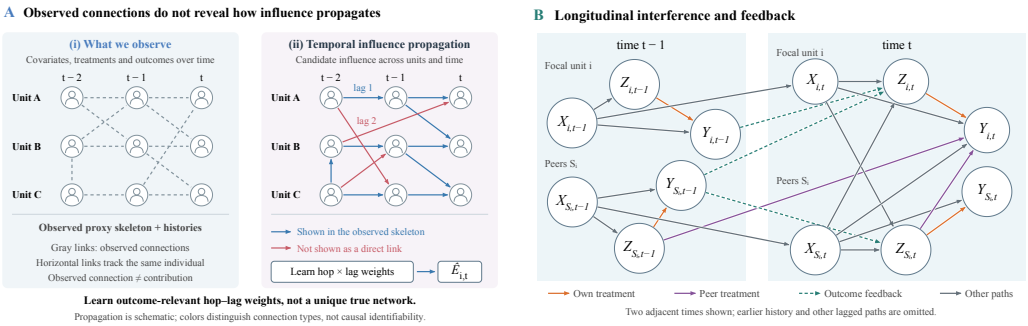


Figure 1: **Dynamic interference from a proxy skeleton and longitudinal histories.** (A) Observed connections specify candidate paths; their outcome-relevant contributions can differ across peers, hops, and lags. DyNICL constructs exposure by learning weights over these paths. (B) Own-treatment paths (orange), lagged peer-treatment paths (purple), and outcome-to-treatment feedback (dashed teal) link adjacent times. Gray arrows denote other dependencies. Peer quantities indexed by S_i collect those units' variables; earlier histories and additional lags are omitted.

2 Related Work

Longitudinal causal estimation. G-methods address time-varying confounding and feedback (Robins, 1986; Robins et al., 1994; Robins, 2004). Robust longitudinal estimation (Laan & Robins, 2003; Van der Laan et al., 2011; Kennedy, 2016; Luedtke & van der Laan, 2016) and off-policy evaluation (Thomas & Brunskill, 2016; Jiang & Li, 2016; Kallus & Uehara, 2020a;b) combine regression and weighting; DyNICL uses them for learned exposure and partial assignment policies.

Interference and learned exposure. Interference research defines direct and spillover effects through assignment restrictions and exposure mappings (Manski, 2013; Tchetgen & VanderWeele, 2012; Benjamin-Chung et al., 2018; Forastiere et al., 2021). Network randomization and rollout designs address dependence in treatment allocation (Ugander et al., 2013; Eckles et al., 2017; Cortez et al., 2022). Representation methods exploit graph information to adjust for covariate imbalance (Veitch et al., 2019; Guo et al., 2020; Cai et al., 2023; Huang et al., 2023). NetEst (Jiang & Sun, 2022) uses learned representations and TNet (Chen et al., 2024) uses targeted learning for network effect estimation. Structure-learning and flexible-aggregation methods relax a fixed interference specification (Bhattacharya et al., 2020; Lin et al., 2024). CauGramer (Wu et al., 2025) learns interference through higher-order aggregation and cross-attention; DyNICL organizes exposure by hop and lag for longitudinal partial-policy recursion. Shared-state interference also motivates population-level inference under Markov and mixing conditions (Hays & Raghavan, 2025). DyNICL instead uses independent network trajectories for cross-fitting.

3 Problem Setup

3.1 Observed Histories and Induced Exposure

We observe R independent network trajectories $O_r = (\mathbf{A}_{\text{obs}}, \mathbf{X}_{r,0:T-1}, \mathbf{Z}_{r,0:T-1}, \mathbf{Y}_{r,0:T-1})$, with units $i = 1, \dots, N$, times $s = 0, \dots, T-1$, and binary treatments $Z_{r,i,s}$. All expectations and limits condition on a fixed design $\mathcal{D}$: the proxy skeleton, reference propagation, structural parameters, pre-generated exogenous covariates, and pre-selected focal/actionable sets. Only observed components enter the learner; treatment draws and outcome innovations vary across the R independent network episodes. We suppress $\mathcal{D}$ in conditional kernels. The pre-assignment history $H_{r,i,s}$ contains network-wide covariates through s and assignments and outcomes strictly before s. The focal index i specifies the intervention; within-step ordering is $H_{r,i,s} \to \mathbf{Z}_{r,s} \to \mathbf{Y}_{r,s}$.

For a finite interference set S_i and memory δ, a reference map summarizes lagged peer assignments:

$$E^{\star}_{r,i,s} = \phi^{\star}_i(\mathbf{A}_{\text{obs}}, \mathbf{Z}_{r,S_i,s-1}, \dots, \mathbf{Z}_{r,S_i,s-\delta}) \,. \tag{1}$$

Exposure sufficiency requires this summary, remaining history, and current actionable assignment to retain the outcome and continuation-state laws. The reference memory δ, policy window L_g, and model lag $\tau_{\max}$ are separate quantities.

3.2 Partial Regimes and Effect Trajectories

Fix a target time t, $s_0 = \max(0, t - L_g)$, and an actionable peer set $S_i^g \subseteq S_i$ selected before fitting or evaluation. Write $\widetilde{S}_i = \{i\} \cup S_i$ and $\mathbf{A}^{\mathrm{loc}}_{r,i,s} = (Z_{r,i,s}, \mathbf{Z}_{r,S_i,s})$. The actionable component $D^g_{r,i,s,t}$ consists of assignments to S_i^g for $s_0 \le s < t$ and the focal treatment $Z_{r,i,t}$ at t. The remaining local assignments form its natural complement $C^g_{r,i,s,t}$. A partial stochastic regime changes only the law of D^g:

$$p^g_{i,s,t}(d, c \mid h) = g_{i,s,t}(d \mid h)\, p_{i,s,t}(c \mid h, d). \tag{2}$$

Assignments outside the local vector also follow their natural conditional law. All transitions use the evolving counterfactual history, including outcome-to-treatment feedback.

Four fixed stochastic regimes g_i^{ab} use peer-policy probability level b before t and own-policy level a at t, with $a, b \in \{0, 1\}$. Policies are independent of fitted propensities; methods share actionable sets and policy levels. Focal and population values are

$$V^{ab}_{i,t} = \mathbb{E}[Y^{g_i^{ab}}_{r,i,t} \mid \mathcal{D}], \tag{3}$$

$$\Psi^{ab}_t = \frac{1}{N} \sum_{i=1}^{N} V^{ab}_{i,t}. \tag{4}$$

Direct, peer, total, and own–peer interaction effects are

$$\mathrm{IDE}^{(0)}_t = \Psi^{10}_t - \Psi^{00}_t, \qquad \mathrm{IPE}^{(0)}_t = \Psi^{01}_t - \Psi^{00}_t, \tag{5}$$

$$\mathrm{ITE}_t = \Psi^{11}_t - \Psi^{00}_t, \qquad \mathrm{INT}_t = \Psi^{11}_t - \Psi^{10}_t - \Psi^{01}_t + \Psi^{00}_t. \tag{6}$$

Hence $\mathrm{ITE}_t = \mathrm{IDE}^{(0)}_t + \mathrm{IPE}^{(0)}_t + \mathrm{INT}_t$. Focal effects use the same contrasts of $V^{ab}_{i,t}$. Each point's policy window ends at its target t. IDE and IPE abbreviate baseline-reference contrasts.

3.3 Identification

Identification requires consistency, sequential exchangeability of $D^g_{i,s,t}$ given $H_{i,s}$, positivity for every regime-supported actionable assignment at regime-reachable histories, exposure sufficiency, and finite interference with well-defined transitions.

After marginalizing the natural complement, define the backward recursion

$$\begin{aligned} q^g_{i,s,t}(h, d) &= \mathbb{E}[\overline{Q}^g_{i,s+1,t}(H_{i,s+1}) \mid H_{i,s} = h, D^g_{i,s,t} = d], \\ \overline{Q}^g_{i,s,t}(h) &= \int q^g_{i,s,t}(h, d) g_{i,s,t}(dd \mid h), \qquad \overline{Q}^g_{i,t+1,t} = Y_{i,t}. \end{aligned} \tag{7}$$

Proposition 3.1 (Identification of regime-value trajectories)**.** *Under the conditions above, every attainable fixed stochastic partial regime g satisfies*

$$\Psi^g_t = \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^{N} \overline{Q}^g_{i,s_0,t}(H_{i,s_0}) \,\middle|\, \mathcal{D} \right]. \tag{8}$$

Appendix B states the assumptions formally and proves identification by backward induction.

4 Methodology

DyNICL first learns exposure and history representations through outcome prediction, then freezes them for policy-specific backward-Q regression (Figure 2). The representation is shared across target times and regimes, while Q-based and augmented estimation use the same fitted continuations.

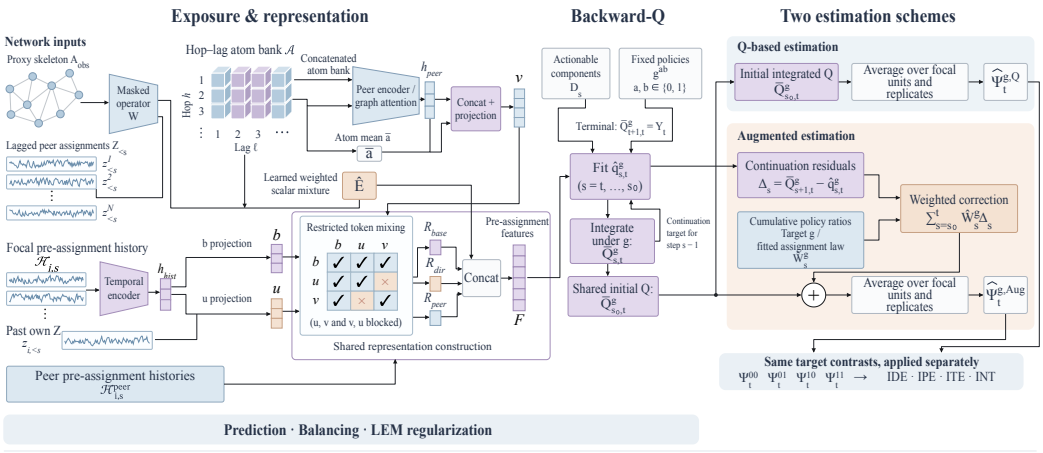


Figure 2: **DyNICL: shared learning, two estimation schemes.** Learned exposure, focal histories, and peer histories feed backward-Q regression and policy integration. Both branches average initial continuations with or without weighted residual correction and apply the same contrasts.

4.1 Learning Dynamic Exposure

The learned exposure module (LEM) preserves spatial reach and temporal delay in a hop–lag bank, with learned edge and lag weights. Let $\mathcal{N}_i^{\text{obs}} = \{j : j \neq i, \mathbf{A}_{\text{obs}}[i,j] > 0\}$ denote one-hop proxy neighbors, distinct from the interference set S_i, and let $L_{ij} = \Theta_{ij} + \lambda_{\text{obs}}\mathbf{A}_{\text{obs}}[i,j]$. For nonempty support, the row-normalized propagation operator is

$$W_\theta[i,j] = \frac{\mathbf{1}\{j \in \mathcal{N}_i^{\text{obs}}\} \exp L_{ij}}{\sum_{u \in \mathcal{N}_i^{\text{obs}}} \exp L_{iu}}. \tag{9}$$

Its powers extend propagation along proxy-supported paths. For positive hops $h = 1, \ldots, H$ and lags $\tau = 1, \ldots, \tau_{\max}$, construct raw atoms and lag-weighted features:

$$B_{i,s}^{(h,\tau)} = \left[(W_\theta^h \odot (\mathbf{1} - I))\mathbf{Z}_{s-\tau}\right]_i, \qquad k_\tau = \text{softmax}(\gamma)_\tau, \tag{10}$$

$$\widehat{E}_{i,s}^{(h,\tau)} = k_\tau B_{i,s}^{(h,\tau)}. \tag{11}$$

The diagonal mask excludes own treatment from each local atom. Positive lags make the atoms available before current assignment. A separate learned mixture supplies scalar exposure:

$$\widehat{E}_{i,s} = \sum_{h=1}^{H}\sum_{\tau=1}^{\tau_{\max}} \alpha_{h,\tau} B_{i,s}^{(h,\tau)}, \qquad \alpha_{h,\tau} = \text{softmax}(\mathbf{a})_{h,\tau}. \tag{12}$$

The encoder receives the full bank; scalar exposure supplies a separate covariate. Both are learned through $\theta = (\Theta, \gamma, \mathbf{a})$. Keeping the atoms separate allows the encoder to distinguish histories with the same scalar exposure but different spatial reach or delay.

4.2 Encoding Histories and Training the Shared Core

To model continuation under feedback, the encoding receives focal and peer covariates through s and treatments and outcomes before s. The temporal encoder produces $\mathbf{h}_{i,s}^{\text{hist}}$ from focal history. The peer encoder uses graph attention (Veličković et al., 2017) biased by $\log W_\theta$, retaining local atoms through a residual projection in $\mathbf{h}_{i,s}^{\text{peer}}$.

Three projected tokens encode context b, own history u, and peer features v. Restricted attention blocks direct u–v attention in both directions while allowing interaction through b, producing $\mathbf{R}^{\text{base}}$, $\mathbf{R}^{\text{dir}}$, and $\mathbf{R}^{\text{peer}}$. The continuation feature is

$$F_{\eta,\theta}(H_{i,s}) = [\mathbf{R}_{i,s}^{\text{base}}, \mathbf{R}_{i,s}^{\text{dir}}, \mathbf{R}_{i,s}^{\text{peer}}, \widehat{E}_{i,s}], \tag{13}$$

where η contains encoder and token-projection parameters.

One-step outcome prediction trains the shared core before continuation fitting. The auxiliary head receives current own treatment after encoding, scalar exposure, and their interaction. Within training fold $\mathcal{T}_{-k}$, optimize

$$\min_{\theta,\eta,\omega,c_E,c_{ZE}} \mathrm{Mean}_{r\in\mathcal{T}_{-k}}\left[\frac{1}{NT}\sum_{i,s}(Y_{r,i,s}-\widehat{Y}_{r,i,s})^2+\lambda_{\mathrm{bal}}\mathcal{L}_{\mathrm{bal},r}+\mathcal{R}_{\mathrm{LEM}}\right]. \tag{14}$$

Here ω, c_E, c_{ZE} parameterize the prediction head; $\mathcal{L}_{\mathrm{bal}}$ aligns treated and untreated base-embedding means (Shalit et al., 2017), and $\mathcal{R}_{\mathrm{LEM}}$ regularizes edge logits and hop–lag mixing entropy.

4.3 Neural Backward-Q and Cross-Fitting

Policy evaluation requires predicting the target outcome as subsequent assignments follow g. We fit these continuations by backward regression, using the shared features learned in Eq. (14). Cross-fitting (Chernozhukov et al., 2018) partitions complete network trajectories into K folds independently of their observations. For each fold k, all models are trained on $\mathcal{T}_{-k}$ and frozen on the held-out fold, preserving within-network dependence.

For each target t and regime g, initialize $\widehat{\overline{Q}}^{g,(-k)}_{i,t+1,t}=Y_{i,t}$ and, for $s=t,\dots,s_0$, regress the next fitted continuation on $\widehat{F}^{(-k)}(H_{i,s})$ and observed actionable assignment $D^g_{i,s,t}$:

$$\widehat{q}^{g,(-k)}_{i,s,t}(h,d)=f_{\widehat{\beta}^{(-k)}_{s,t,g}}(\widehat{F}^{(-k)}(h),d), \tag{15}$$

$$\widehat{\overline{Q}}^{g,(-k)}_{i,s,t}(h)=\int \widehat{q}^{g,(-k)}_{i,s,t}(h,d)g_{i,s,t}(dd\mid h). \tag{16}$$

Integration under g supplies the regression target for step $s-1$; conditioning on history and D^g marginalizes the natural complement. The analysis assumes that the frozen features retain the continuation mean in Eq. (7) for each step and regime, and that the head class can approximate it.

4.4 Two Regime-Value Estimation Schemes

Let $k(r)$ denote replicate r's held-out fold.

Q-based estimation. The plug-in estimate averages the initial integrated continuation:

$$\widehat{\Psi}^{g,\mathrm{Q}}_t=\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{i=1}^{N}\widehat{\overline{Q}}^{g,(-k(r))}_{i,s_0,t}(H_{r,i,s_0}). \tag{17}$$

Augmented estimation. Augmentation adds continuation residuals weighted from the observed law to the target policy (Bang & Robins, 2005; Jiang & Li, 2016). The natural-complement factors cancel in the partial kernel, leaving ratios of actionable-assignment probabilities. For the fitted law $\widehat{p}^{D,(-k)}_{i,s,t}(d\mid h)$, define

$$\widehat{\rho}^{g,(-k)}_{r,i,s,t}=\frac{g_{i,s,t}(D^g_{r,i,s,t}\mid H_{r,i,s})}{\widehat{p}^{D,(-k)}_{i,s,t}(D^g_{r,i,s,t}\mid H_{r,i,s})},\qquad \widehat{W}^{g,(-k)}_{r,i,s,t}=\prod_{u=s_0}^{s}\widehat{\rho}^{g,(-k)}_{r,i,u,t}. \tag{18}$$

For multi-node actions, we use the conditional Bernoulli factorization and strong positivity in Assumption B.4.

The held-out residual, corrected score, and estimator are

$$\widehat{\Delta}^{g,(-k)}_{r,i,s,t}=\widehat{\overline{Q}}^{g,(-k)}_{i,s+1,t}(H_{r,i,s+1})-\widehat{q}^{g,(-k)}_{i,s,t}(H_{r,i,s},D^g_{r,i,s,t}), \tag{19}$$

$$\widehat{\psi}^{g,(-k)}_{r,i,t}=\widehat{\overline{Q}}^{g,(-k)}_{i,s_0,t}(H_{r,i,s_0})+\sum_{s=s_0}^{t}\widehat{W}^{g,(-k)}_{r,i,s,t}\widehat{\Delta}^{g,(-k)}_{r,i,s,t}, \tag{20}$$

$$\widehat{\Psi}^{g,\mathrm{Aug}}_t=\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{i=1}^{N}\widehat{\psi}^{g,(-k(r))}_{r,i,t}. \tag{21}$$

The terminal residual uses $Y_{r,i,t}$. Averaging over replicates at fixed i gives the focal estimates $\widehat{V}_{i,t}^{g,\mathrm{Q}}$ and $\widehat{V}_{i,t}^{g,\mathrm{Aug}}$. Each branch separately applies the four effect contrasts.

4.5 Estimator-Specific Guarantees

For fixed K and nonvanishing fold proportions, square-integrable, conditionally L_2-consistent initial continuations give Q-based consistency with first-order mean-continuation error (Proposition B.19).

For augmented estimation, Theorem B.13 gives

$$\widehat{\Psi}_t^{g,\mathrm{Aug}} - \Psi_t^g = \frac{1}{R}\sum_{r=1}^{R} \varphi_t^g(O_r) + O_p(\mathfrak{R}_{q,p,t}) + O_p(\epsilon_\phi) + o_p(R^{-1/2}), \tag{22}$$

where φ_t^g is the centered true replicate score, R_k the fold size, (q_0, p_0) the true nuisances, and

$$\mathfrak{R}_{q,p,t} = \sum_{k=1}^{K} \frac{R_k}{R} \|\widehat{q}^{(-k)} - q_0\|_{\mathcal{Q},t} \|\widehat{p}^{(-k)} - p_0\|_{\mathcal{W},t}.$$

The norms aggregate continuation and cumulative-ratio errors across time. At reference exposure, the compatible score is doubly robust and Neyman-orthogonal in (q, p), with $O_p(\epsilon_\phi)$ exposure drift.

The expansion holds under the identification conditions, independent replicates, exact compatible integration, uncapped weights, nonvanishing fold proportions, square-integrable true scores, fitted-score L_2 convergence, and fold-conditional exposure stability (Assumption B.11), as $R \to \infty$ with network size, horizon, regimes, and K fixed. Root-R inference additionally requires both remainders to be $o_p(R^{-1/2})$.

Appendices B–C provide proofs, numerical-integration and weight-capping results, and implementation details.

5 Experimental Evaluation

5.1 Evaluation Setup

Data and targets. The synthetic grid crosses expected neighbor count $d \in \{4, 8, 12\}$ with generating lag $L \in \{1, 3, 5\}$ under hop–lag exposure and treatment–exposure interaction. Propagation weights are fixed within each design, while assignments, outcomes, and exposure evolve over time. Each run has 30 independent trajectories. Methods share data, folds, focal units, evaluation times, and four stochastic policies with window L. Each synthetic trajectory contains 64 nodes over 12 time steps. Own-policy probabilities are $0.2/0.8$ and peer-policy probabilities are $0.3/0.7$, with at most three actionable peers per focal unit.

Comparators. We compare DyNICL-Q and DyNICL-Aug with NetEst (Jiang & Sun, 2022), TNet (Chen et al., 2024), and three exposure baselines: fixed one-hop, fixed multi-hop, and ObsGraph. Each exposure baseline uses both Q-based and augmented estimation. Fixed one-hop averages immediate neighbors at lag one; fixed multi-hop aggregates one to three hops over the setting's lag window. ObsGraph fixes propagation weights to the observed proxy graph and learns how its hop–lag exposure features are combined.

Metrics. We aggregate focal-by-time RMSE over 10 independent runs as mean $\pm$ sample standard deviation; lower is better. Tables 1–2 show means; plot error bars span one standard deviation. Appendices D–E provide the experimental protocols and complete results for all four effects, including the full sensitivity and robustness plots.

5.2 Comparison across Neighborhoods and Memory

DyNICL-Q has lower IPE and ITE RMSE than NetEst and TNet across all nine settings (Table 1).

Table 1: Mean IPE and ITE RMSE over d neighbors and generating lag L. Bold marks the lowest mean RMSE for each setting and effect.

Method	4×1	4×3	4×5	8×1	8×3	8×5	12×1	12×3	12×5
IPE									
Fixed one-hop-Q	0.264	0.436	0.664	0.427	0.673	0.963	0.783	1.048	1.338
Fixed one-hop-Aug	0.247	0.386	0.758	0.365	0.597	1.136	0.571	1.013	1.689
Fixed multi-hop-Q	0.228	0.292	0.354	0.363	0.449	0.558	0.689	0.796	0.943
Fixed multi-hop-Aug	0.231	0.365	0.673	0.334	0.538	0.947	0.509	0.932	1.454
ObsGraph-Q	0.221	0.268	0.318	0.342	0.404	0.483	0.633	0.715	0.826
ObsGraph-Aug	0.224	0.332	0.594	0.317	0.482	0.833	0.453	0.822	1.285
NetEst	0.354	0.407	0.459	0.442	0.490	0.608	0.684	0.721	0.846
TNet	0.343	0.411	0.443	0.428	0.474	0.593	0.667	0.703	0.825
DyNICL-Q	**0.209**	**0.247**	**0.286**	**0.303**	**0.355**	**0.406**	0.530	**0.586**	**0.676**
DyNICL-Aug	0.218	0.316	0.553	0.308	0.438	0.762	**0.387**	0.764	1.174
ITE									
Fixed one-hop-Q	0.402	0.588	0.857	0.649	0.900	1.275	1.179	1.382	1.746
Fixed one-hop-Aug	0.379	0.532	0.967	0.572	0.798	1.527	0.817	1.279	2.219
Fixed multi-hop-Q	0.363	0.421	0.464	0.562	0.622	0.742	1.067	1.104	1.278
Fixed multi-hop-Aug	0.359	0.483	0.790	0.526	0.703	1.273	0.739	1.149	1.930
ObsGraph-Q	0.357	0.389	0.419	0.538	0.559	0.633	0.987	0.979	1.097
ObsGraph-Aug	0.352	0.444	0.701	0.502	0.622	1.109	0.673	0.987	1.693
NetEst	0.553	0.665	0.693	0.615	0.664	0.775	1.095	1.098	1.143
TNet	0.535	0.647	0.680	0.597	0.646	0.753	1.071	1.100	1.116
DyNICL-Q	**0.345**	**0.365**	**0.382**	0.486	**0.498**	**0.519**	0.876	**0.839**	**0.920**
DyNICL-Aug	0.348	0.423	0.650	**0.475**	0.563	0.998	**0.590**	0.912	1.528

Table 2: Module ablations. Entries give mean RMSE as Q/Aug, sharing the same ablated core. Columns identify $d \times L$.

Variant	Effect	4×1	4×5	8×3	12×1	12×5
DyNICL	IPE	0.209/0.218	0.286/0.553	0.355/0.438	0.530/0.387	0.676/1.174
	ITE	0.345/0.348	0.382/0.650	0.498/0.563	0.876/0.590	0.920/1.528
	INT	0.156/0.173	0.227/0.449	0.284/0.368	0.421/0.433	0.493/0.948
w/o peer pathway	IPE	0.369/0.317	0.578/0.714	0.730/0.634	0.984/0.684	1.328/1.729
	ITE	0.506/0.454	0.716/0.892	0.917/0.809	1.398/0.975	1.747/2.267
	INT	0.285/0.258	0.492/0.633	0.615/0.550	0.827/0.708	1.099/1.469
w/o temporal attention	IPE	0.212/0.221	0.322/0.589	0.384/0.466	0.536/0.392	0.743/1.262
	ITE	0.350/0.353	0.537/0.772	0.586/0.643	0.885/0.595	1.163/1.794
	INT	0.159/0.177	0.339/0.544	0.345/0.424	0.427/0.438	0.682/1.153
w/o peer attention	IPE	0.225/0.230	0.309/0.578	0.407/0.476	0.682/0.491	0.883/1.382
	ITE	0.362/0.361	0.407/0.679	0.557/0.612	1.064/0.732	1.198/1.846
	INT	0.169/0.185	0.246/0.471	0.329/0.410	0.549/0.539	0.703/1.193
w/o token mixer	IPE	0.214/0.223	0.303/0.572	0.371/0.449	0.549/0.400	0.781/1.279
	ITE	0.346/0.351	0.420/0.691	0.530/0.587	0.899/0.607	1.125/1.743
	INT	0.159/0.174	0.254/0.480	0.308/0.388	0.438/0.447	0.659/1.122

With the same Q-based estimator and hop range, learned rather than fixed multi-hop aggregation reduces ITE RMSE from 1.2783 to 0.9197 at 12×5. DyNICL also improves over ObsGraph and fixed one-hop within each branch; the latter gap is largest beyond one-step memory.

With shared DyNICL continuations, augmentation reduces ITE from 0.8762 to 0.5896 at 12×1. At 12×5, its ITE error is instead 1.5284, above Q-based estimation (0.9197) and TNet (1.1160).

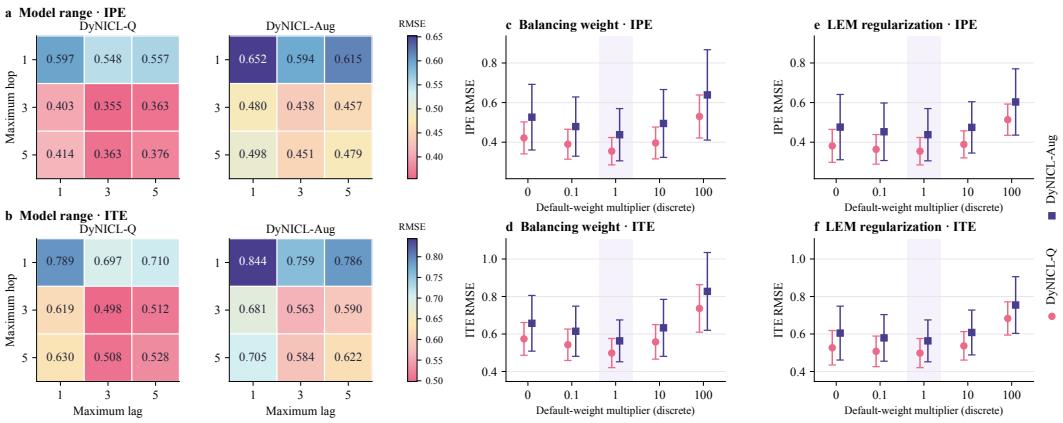


Figure 3: **Model-range and regularization sensitivity on 8×3 data.** (a,b) IPE and ITE errors over the model's maximum hop and lag; each Q/Aug heatmap pair shares a color scale, with pink indicating lower error and purple higher error. (c,d) Balancing-weight and (e,f) LEM-regularization sweeps use discrete multiples of the default coefficient. Pale shading marks the lowest observed mean in each weight sweep.

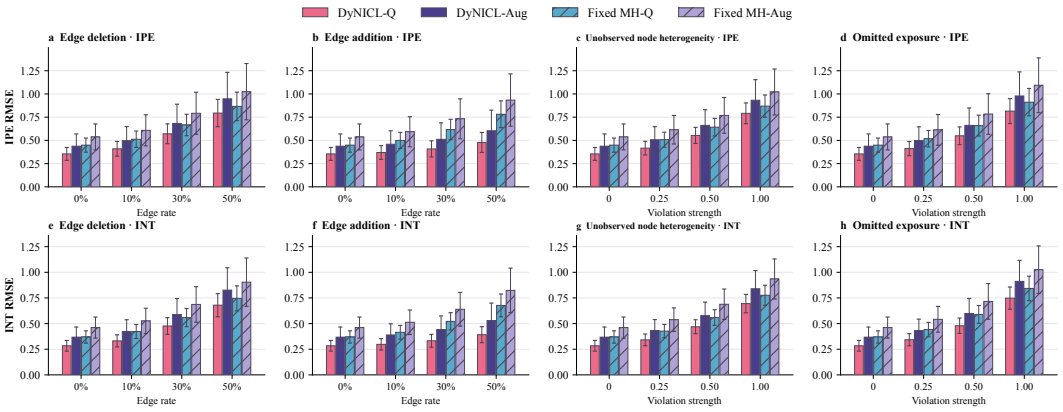


Figure 4: **Robustness to graph error and omitted information.** IPE (top) and INT (bottom) on 8×3 data; hatched bars denote fixed multi-hop. Edge rates use the original edge count, holding data and targets fixed. Node effects and omitted exposure change both the SCM and causal truth.

5.3 Contributions of the Shared Modules

Table 2 isolates four modules at the grid corners and center, holding policies fixed. Deleting the full peer pathway, including LEM and scalar exposure, causes the largest IPE increase in every setting. Even when local LEM features remain available, removing peer attention raises Q-based IPE from 0.6763 to 0.8834 at 12×5.

Temporal attention contributes more under long memory: its removal changes Q-based ITE from 0.3448 to 0.3496 at 4×1, but from 0.3821 to 0.5369 at 4×5, despite retaining lagged exposure. Removing token mixing raises INT from 0.4931 to 0.6593 at 12×5, although it slightly improves IDE at 4×1 (0.2293 to 0.2271).

5.4 Sensitivity and Robustness

Model range and regularization. At fixed 8×3 data and policy window (Figure 3), Q-based IPE and ITE errors are lowest at $(H, \tau_{\max}) = (3, 3)$ among matched ranges $(1, 1)$, $(3, 3)$, and $(5, 5)$.

Table 3: SocioPatterns semi-synthetic comparison. Contact structure is observed; treatments, outcomes, and effects are simulated. Cases differ in both memory and feedback. Bold marks the lowest mean per case and effect.

Method	IDE	IPE	ITE	INT
Short memory ($L = 1$), no feedback				
DyNICL-Q	$\mathbf{0.274}_{\pm 0.063}$	$0.324_{\pm 0.064}$	$0.570_{\pm 0.066}$	$\mathbf{0.315}_{\pm 0.055}$
DyNICL-Aug	$0.382_{\pm 0.027}$	$\mathbf{0.302}_{\pm 0.025}$	$\mathbf{0.504}_{\pm 0.071}$	$0.372_{\pm 0.061}$
Fixed multi-hop-Q	$0.375_{\pm 0.065}$	$0.422_{\pm 0.074}$	$0.653_{\pm 0.070}$	$0.417_{\pm 0.067}$
Fixed multi-hop-Aug	$0.496_{\pm 0.026}$	$0.420_{\pm 0.053}$	$0.614_{\pm 0.092}$	$0.590_{\pm 0.062}$
TNet	$0.502_{\pm 0.083}$	$0.431_{\pm 0.078}$	$0.665_{\pm 0.146}$	$0.478_{\pm 0.072}$
NetEst	$0.527_{\pm 0.119}$	$0.440_{\pm 0.100}$	$0.676_{\pm 0.131}$	$0.493_{\pm 0.077}$
Long memory ($L = 5$) with feedback				
DyNICL-Q	$\mathbf{0.243}_{\pm 0.040}$	$0.418_{\pm 0.170}$	$1.009_{\pm 0.215}$	$0.389_{\pm 0.074}$
DyNICL-Aug	$0.310_{\pm 0.032}$	$\mathbf{0.309}_{\pm 0.194}$	$\mathbf{0.781}_{\pm 0.374}$	$\mathbf{0.225}_{\pm 0.097}$
Fixed multi-hop-Q	$0.343_{\pm 0.038}$	$0.812_{\pm 0.166}$	$1.294_{\pm 0.209}$	$0.598_{\pm 0.083}$
Fixed multi-hop-Aug	$0.479_{\pm 0.127}$	$0.791_{\pm 0.183}$	$1.277_{\pm 0.323}$	$0.592_{\pm 0.055}$
TNet	$0.548_{\pm 0.142}$	$0.724_{\pm 0.175}$	$1.143_{\pm 0.218}$	$0.451_{\pm 0.073}$
NetEst	$0.619_{\pm 0.171}$	$0.663_{\pm 0.192}$	$1.002_{\pm 0.320}$	$0.309_{\pm 0.067}$

Varying the two axes separately shows that temporal coverage alone does not recover the benefit of spatial coverage: at maximum lag three, increasing the hop range from one to three lowers Q-based IPE from 0.5483 to 0.3546. Extending both ranges to five raises it to 0.3764. Removing either regularizer or using 100 times its default weight increases error.

Graph error and omitted information. DyNICL retains lower IPE and INT errors than fixed multi-hop under all four perturbations (Figure 4). Both methods deteriorate with graph error and omitted information; deletion removes candidate paths and is particularly damaging.

Assignment support. On 8×3 and 12×5, augmentation has slightly lower errors at generating probability floor $\epsilon = 0.20$, but its error and between-run dispersion increase more sharply as the floor decreases. Estimated propensities remain clipped to $[0.01, 0.99]$ throughout.

5.5 Semi-Synthetic Contact-Network Study

Using the 180-node SocioPatterns contact network (Fournet & Barrat, 2014), we simulate treatments and outcomes under the same four policies in two settings: short memory ($L = 1$) without feedback and long memory ($L = 5$) with feedback.

DyNICL-Q has the lowest IDE error in both cases, and DyNICL-Aug has the lowest IPE and ITE errors (Table 3). Under long memory and feedback, augmented IPE falls from 0.791 for fixed multi-hop to 0.309 for DyNICL. INT favors Q for short memory and Aug for long memory with feedback. The trade-off is effect-specific within the same dataset: in the long-memory case, augmentation lowers DyNICL's ITE from 1.009 to 0.781 while increasing IDE from 0.243 to 0.310.

6 Conclusion

DyNICL couples learned hop–lag exposure with neural backward-Q to estimate four dynamic effect trajectories. The analysis establishes Q-based consistency and an augmented expansion with product-error and exposure-drift terms. Experiments show lower peer and total errors than external baselines, with Q-based estimation stronger in most main-grid settings and augmentation stronger for peer and total effects in both semi-synthetic cases. These guarantees require independent network trajectories, representation sufficiency, and nuisance convergence. Empirical evidence is limited to synthetic and semi-synthetic outcomes; augmentation remains sensitive to weak assignment support.

AI USE STATEMENT

During manuscript preparation, AI tools were used for language polishing and improving readability. The authors reviewed and edited the text and take full responsibility for the final manuscript.

REPRODUCIBILITY STATEMENT

Appendix B provides the assumptions and proofs for identification and the two estimation schemes. Appendix C details the model architecture, training objectives, nuisance estimation, and default configurations. Appendix D specifies the data-generating process, policy interventions, oracle evaluation, baseline adaptations, semi-synthetic data construction, and evaluation protocols. Appendix E reports the full synthetic results, including variability across independent runs, module ablations, and sensitivity and robustness analyses.

REFERENCES

Peter M. Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4), December 2017. ISSN 1932-6157. doi: 10.1214/16-aoas1005. URL http://dx.doi.org/10.1214/16-AOAS1005.

Heejung Bang and James M Robins. Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4):962–973, 2005.

Jade Benjamin-Chung, Benjamin F Arnold, David Berger, Stephen P Luby, Edward Miguel, John M Colford Jr, and Alan E Hubbard. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *International journal of epidemiology*, 47(1):332–347, 2018.

Rohit Bhattacharya, Daniel Malinsky, and Ilya Shpitser. Causal inference under interference and network uncertainty. In *Uncertainty in Artificial Intelligence*, pp. 1028–1038. PMLR, 2020.

Ruichu Cai, Zeqin Yang, Weilin Chen, Yuguang Yan, and Zhifeng Hao. Generalization bound for estimating causal effects from observational network data. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pp. 163–172, 2023.

Weilin Chen, Ruichu Cai, Zeqin Yang, Jie Qiao, Yuguang Yan, Zijian Li, and Zhifeng Hao. Doubly robust causal effect estimation under networked interference via targeted learning. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 6457–6485. PMLR, 2024. URL https://proceedings.mlr.press/v235/chen24c.html.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters, 2018.

Mayleen Cortez, Matthew Eichhorn, and Christina Yu. Staggered rollout designs enable causal inference under interference without network knowledge. *Advances in Neural Information Processing Systems*, 35:7437–7449, 2022.

Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1):20150021, 2017.

Laura Forastiere, Edoardo M Airoldi, and Fabrizia Mealli. Identification and estimation of treatment and interference effects in observational studies on networks. *Journal of the American Statistical Association*, 116(534):901–918, 2021.

Laura Forastiere, Fabrizia Mealli, Albert Wu, and Edoardo M Airoldi. Estimating causal effects under network interference with bayesian generalized propensity scores. *Journal of Machine Learning Research*, 23(289):1–61, 2022.

Julie Fournet and Alain Barrat. Contact patterns among high school students. *PloS one*, 9(9): e107878, 2014.

Ruocheng Guo, Jundong Li, and Huan Liu. Learning individual causal effects from networked observational data. In *Proceedings of the 13th international conference on web search and data mining*, pp. 232–240, 2020.

Chris Hays and Manish Raghavan. Double machine learning for causal inference under shared-state interference. *arXiv preprint arXiv:2504.08836*, 2025.

Miguel A Hernán and James M Robins. Causal inference, 2010.

Qiang Huang, Jing Ma, Jundong Li, Ruocheng Guo, Huiyan Sun, and Yi Chang. Modeling interference for individual treatment effect estimation from networked observational data. *ACM Transactions on Knowledge Discovery from Data*, 18(3):1–21, 2023.

Michael G Hudgens and M Elizabeth Halloran. Toward causal inference with interference. *Journal of the american statistical association*, 103(482):832–842, 2008.

Nan Jiang and Lihong Li. Doubly robust off-policy value evaluation for reinforcement learning. In *International conference on machine learning*, pp. 652–661. PMLR, 2016.

Song Jiang and Yizhou Sun. Estimating causal effects on networked observational data via representation learning. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 852–861, 2022.

Nathan Kallus and Masatoshi Uehara. Double reinforcement learning for efficient off-policy evaluation in markov decision processes. *Journal of Machine Learning Research*, 21(167):1–63, 2020a.

Nathan Kallus and Masatoshi Uehara. Double reinforcement learning for efficient and robust off-policy evaluation. In *International Conference on Machine Learning*, pp. 5078–5088. PMLR, 2020b.

Edward H Kennedy. Semiparametric theory and empirical processes in causal inference. In *Statistical causal inferences and their applications in public health research*, pp. 141–167. Springer, 2016.

Mark J Laan and James M Robins. *Unified methods for censored longitudinal data and causality*. Springer, 2003.

Xiaofeng Lin, Guoxi Zhang, Xiaotian Lu, and Hisashi Kashima. Treatment effect estimation under unknown interference. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 28–42. Springer, 2024.

Alexander R Luedtke and Mark J van der Laan. Super-learning of an optimal dynamic treatment rule. *The international journal of biostatistics*, 12(1):305–332, 2016.

Charles F Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013.

James Robins. A new approach to causal inference in mortality studies with a sustained exposure periodapplication to control of the healthy worker survivor effect. *Mathematical modelling*, 7 (9-12):1393–1512, 1986.

James M Robins. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium in Biostatistics: analysis of correlated data*, pp. 189–326. Springer, 2004.

James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, 89 (427):846–866, 1994.

Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pp. 3076–3085. PMLR, 2017.

Eric J Tchetgen Tchetgen and Tyler J VanderWeele. On causal inference in the presence of interference. *Statistical methods in medical research*, 21(1):55–75, 2012.

Philip Thomas and Emma Brunskill. Data-efficient off-policy policy evaluation for reinforcement learning. In *International conference on machine learning*, pp. 2139–2148. PMLR, 2016.

Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 329–337, 2013.

Mark J Van der Laan, Sherri Rose, et al. *Targeted learning: causal inference for observational and experimental data*, volume 4. Springer, 2011.

Victor Veitch, Yixin Wang, and David Blei. Using embeddings to correct for unobserved confounding in networks. *Advances in Neural Information Processing Systems*, 32, 2019.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Anpeng Wu, Haiyi Qiu, Zhengming Chen, Zijian Li, Ruoxuan Xiong, Fei Wu, and Kun Zhang. Causal graph transformer for treatment effect estimation under unknown interference. In *The Thirteenth International Conference on Learning Representations*, 2025. URL `https://openreview.net/forum?id=foQ4AeEGG7`.

A NOTATION

Indices and observed data. $r \in \{1, \dots, R\}$ indexes independent network trajectories, $i \in \{1, \dots, N\}$ indexes units, and $s, t \in \{0, \dots, T-1\}$ index time. The observed variables are covariates $\mathbf{X}_{r,i,s}$, binary assignments $Z_{r,i,s}$, outcomes $Y_{r,i,s}$, and the assignment vector $\mathbf{Z}_{r,s}$.

Design and history. $\mathcal{D}$ is the fixed network design indexing the conditional trajectory law $P_{\mathcal{D}}$ (Section 3.1). $\mathbf{A}_{\text{obs}}$ is the fixed proxy skeleton of candidate ties. $\mathcal{F}^-_{r,s}$ is the pre-assignment filtration, and $H_{r,i,s}$ denotes the full pre-assignment history used for identification. The within-step order is $\mathcal{F}^-_{r,s} \to \mathbf{Z}_{r,s} \to \mathbf{Y}_{r,s}$.

Local interference and exposure. S_i is the focal interference set, $S_i^g \subseteq S_i$ is the pre-specified actionable peer set, and $\mathbf{A}^{\text{loc}}_{r,i,s} = (Z_{r,i,s}, \mathbf{Z}_{r,S_i,s})$ is the physical local assignment vector. For target t, $I^g_{i,s,t} = S_i^g$ before t and $I^g_{i,t,t} = \{i\}$; $D^g_{r,i,s,t} = \mathbf{Z}_{r,I^g_{i,s,t},s}$ is the actionable assignment and the complement remains natural. The reference map $\phi_i^\star$ summarizes past peer assignments over lags $1, \dots, \delta$ as exposure $E^\star_{r,i,s}$; $\widehat{E}^{(-k)}_{r,i,s}$ is its fold-specific learned approximation. W_θ and the hop–lag weights $\alpha_{h,\tau}$ parameterize the DyNICL exposure module, with hop limit H and lag limit $\tau_{\max}$. The policy window L_g sets $s_0 = \max(0, t - L_g)$ and is distinct from the reference memory δ and learner lag.

Regimes and effects. g_i^{ab}, $a, b \in \{0, 1\}$, denotes a fixed stochastic four-cell regime on own and actionable-peer assignments. Its focal value is $V^{ab}_{i,t}$, and its network-average target is $\Psi^{ab}_t = \mathbb{E}[N^{-1} \sum_i Y^{g_i^{ab}}_{r,i,t} \mid \mathcal{D}]$. IDE, IPE, ITE, and INT are the regime-value contrasts defined in Section 3.2.

Prediction and continuation nuisances. m_ω is the single-step prediction head. For regime g, $q^g_{i,s,t}$ is the backward continuation regression and $\overline{Q}^g_{i,s,t}$ is its regime-integrated value. These functions are distinct from the exposure map ϕ.

Orthogonal estimation. $\nu = (q, p)$ denotes the causal nuisance collection and is distinct from the encoder parameter η. $p^D_{i,s,t}$ is the actionable-assignment law, $\rho^g_{i,s,t} = g_{i,s,t}/p^D_{i,s,t}$ is the step regime ratio, and $W^g_{i,s,t}$ is its cumulative product. The held-out augmented score is $\psi^g_{r,i,t}$ and the cross-fitted regime-value estimate is $\widehat{\Psi}^g_t$ (or $\widetilde{\Psi}^g_{M,t}$ with M Monte Carlo regime draws). The term ϵ_ϕ denotes exposure-learning drift. DyNICL provides Q-based and augmented estimates, denoted $\widehat{\Psi}^{g,\text{Q}}_t$ and $\widehat{\Psi}^{g,\text{Aug}}_t$. The unqualified score notation in the theoretical appendix refers to augmented estimation.

B IDENTIFICATION AND ESTIMATION THEORY

B.1 FILTRATION, LOCAL INTERVENTIONS, AND ASSUMPTIONS

Fix the design $\mathcal{D}$ defined in Section 3.1. Throughout this appendix, $P = P_{\mathcal{D}}$, expectations, L_2 norms, and stochastic orders refer to this conditional trajectory law. Structural parameters index P without being inputs to the fitted models. Training fold conditioning additionally fixes that fold's observations and fitting randomness. Let $r = 1, \dots, R$ index independent network trajectories and let

$$\mathcal{F}^-_{r,s} = \sigma\{\mathbf{A}_{\mathrm{obs}}, \mathbf{X}_{r,0:s}, \mathbf{Z}_{r,0:s-1}, \mathbf{Y}_{r,0:s-1}\}$$

be the pre-assignment filtration. For a policy window L_g, write $s_0 = \max(0, t - L_g)$; the reference exposure memory δ need not equal L_g or the learner lag $\tau_{\max}$. The temporal order is $\mathcal{F}^-_{r,s} \to \mathbf{Z}_{r,s} \to \mathbf{Y}_{r,s} \to \mathcal{F}^-_{r,s+1}$. For focal unit i, S_i is its finite interference set and $\mathbf{A}^{\mathrm{loc}}_{r,i,s}$ is the assignment vector on $\{i\} \cup S_i$. For target t, the regime-specific actionable set is $I^g_{i,s,t} = S^g_i$ for $s < t$ and $I^g_{i,t,t} = \{i\}$. Write

$$D^g_{r,i,s,t} = \mathbf{Z}_{r,I^g_{i,s,t},s}, \qquad C^g_{r,i,s,t} = \mathbf{Z}_{r,\widetilde{S}_i \setminus I^g_{i,s,t},s}$$

for the actionable and natural-complement parts of the local vector. The partial intervention kernel is

$$p^g_{i,s,t}(d, c \mid h) = g_{i,s,t}(d \mid h) p_{i,s,t}(c \mid h, d). \tag{23}$$

It replaces only the actionable law, including along counterfactual histories; the complement retains its observed conditional transition. Exposure is the induced summary

$$E^\star_{r,i,s} = \phi^\star_i(\mathbf{A}_{\mathrm{obs}}, \mathbf{Z}_{r,S_i,s-1}, \dots, \mathbf{Z}_{r,S_i,s-\delta})$$

under the intervened assignment history.

For a realized actionable-assignment prefix $\bar{d}_{s_0:s-1} = (d_{s_0}, \dots, d_{s-1})$, let $H^{\bar{d}_{s_0:s-1}}_{r,i,s}$ denote the history generated by applying that prefix while drawing each complement from the natural kernel in Eq. (23). Let $\mathcal{C}^{\bar{d}_{s_0:s-1}, g_{s:t}}_{r,i,s:t}$ collect the subsequent counterfactual histories and terminal outcome obtained by applying the continuation regime $g_{s:t}$. Let $P^g_{H_s}$ denote the history law induced by a partial regime through time $s-1$. Throughout, statements at *regime-reachable histories* hold $P^g_{H_s}$-almost everywhere, and analogously for supported partial assignment prefixes. All spaces are standard Borel, so regular conditional kernels exist; conditional equalities are understood up to the corresponding null sets. Throughout the identification argument, $H_{r,i,s}$ denotes the full pre-assignment history, including earlier actionable assignments and the information needed for all subsequent conditional kernels. Thus cumulative ratios through time $s-1$ are measurable with respect to $H_{r,i,s}$. The focal index specifies the intervention. Write $H_{r,i,s} = (B_{r,i,s}, L_{r,i,s})$, where $L_{r,i,s}$ is the finite past peer-assignment segment summarized by $\phi^\star$, and $B_{r,i,s}$ contains the remaining pre-assignment history, including focal and peer outcome histories and the covariates needed for their natural transitions. The exposure-sufficient state is $R^\star_{r,i,s} = (B_{r,i,s}, E^\star_{r,i,s})$. A neural representation compresses it.

Assumption B.1 (Consistency)**.** If the observed actionable prefix equals $\bar{d}_{s_0:s}$, and the corresponding complement follows its observed natural draw, then the observed history through $s+1$ equals $H^{\bar{d}_{s_0:s}}_{r,i,s+1}$. If the complete observed actionable path agrees with regime g_i from s_0 through t, then $Y_{r,i,t} = Y^{g_i}_{r,i,t}$.

Assumption B.2 (Sequential exchangeability)**.** For every $s \in \{s_0, \dots, t\}$, every assignment prefix with a regime-reachable history, and every fixed continuation regime $g_{s:t}$, the observed actionable assignment is conditionally independent of the entire future counterfactual continuation:

$$\mathcal{C}^{\bar{d}_{s_0:s-1}, g_{s:t}}_{r,i,s:t} \perp\!\!\!\perp D^g_{r,i,s,t} \mid H_{r,i,s},$$

on the event that the prior assignment prefix is $\bar{d}_{s_0:s-1}$. Here $H_{r,i,s}$ is measurable with respect to $\mathcal{F}^-_{r,s}$. Sequential randomization covers both the next state and terminal outcome under intervention.

Assumption B.3 (Assignment-space positivity)**.** For every target regime and time, for $P^g_{H_s}$-almost every h, the actionable policy is absolutely continuous with respect to the observed actionable-assignment kernel:

$$g_{i,s,t}(\cdot \mid h) \ll p^D_{i,s,t}(\cdot \mid h).$$

For the finite binary assignment vectors considered here, this means

$$g_{i,s,t}(d \mid h) > 0 \quad \Longrightarrow \quad p^D_{i,s,t}(d \mid h) > 0.$$

The cumulative regime ratios have a finite $2 + \epsilon$ moment for some $\epsilon > 0$.

Assumption B.4 (Assignment factorization and strong positivity)**.** For the assignment nuisance used in the implementation, conditionally on the full pre-assignment history,

$$p^D_{i,s,t}(d \mid h) = \prod_{j \in I^g_{i,s,t}} \pi_{j,s}(h)^{d_j} \{1 - \pi_{j,s}(h)\}^{1-d_j}, \qquad \kappa \le \pi_{j,s}(h) \le 1 - \kappa$$

for a fixed $\kappa > 0$, uniformly over actionable nodes, times, and regime-reachable histories. The fixed stochastic regime has the corresponding explicit nodewise factorization. The implementation projects fitted probabilities to $[c, 1 - c]$ with a fixed $0 < c < \kappa$. If the unprojected estimator is consistent, this projection preserves its probability limit. The exact orthogonal estimator has no cumulative-weight cap. A capped implementation is analyzed separately in Proposition B.15.

Assumption B.5 (Exposure sufficiency)**.** Conditional on the same $B_{r,i,s}$ and actionable assignment, two values of $L_{r,i,s}$ that produce the same $E^\star_{r,i,s}$ induce the same conditional law of the focal outcome and of the next state required by the target continuation value. Equivalently, after the natural complement is marginalized, the relevant observed outcome and transition kernels factor through $R^\star_{r,i,s}$ and the actionable assignment.

Assumption B.6 (Finite interference and regime transition)**.** Assignments outside $\{i\} \cup S_i$ affect the focal process only through variables already contained in $H_{r,i,s}$. The conditional state-transition law remains well defined under every target regime.

Assumption B.7 (Independent network trajectories)**.** Conditional on $\mathcal{D}$, $O_1, \ldots, O_R$ are independent and identically distributed network trajectories, with no cross-replicate interference.

Representation and approximation conditions. Let $F_{\eta,\phi}(H_s)$ denote the actual pre-assignment features supplied to a continuation head, formed from focal and peer histories together with exposure features as in Eq. (47). Interpreting neural fits as consistent estimates of $q_{0,s}$ requires that these features retain the relevant continuation mean:

$$q_{0,s}(H_s, d) = f_{0,s}\{F_{\eta,\phi^\star}(H_s), d\} \quad \text{almost surely}$$

for a function in the closure of the fitted class. Assignment estimation requires its own sufficient feature set. If either representation or function class is misspecified, its approximation error remains in the corresponding q or p norm. The expansion additionally requires fitted-score convergence and exposure stability; the latter controls changes through T_ϕ at fixed nuisance functions.

B.2 Recursive identification

For $s_0 = \max(0, t - L_g)$, set $\overline{Q}^g_{i,t+1,t} = Y_{i,t}$ and recursively define, for $s = t, \ldots, s_0$,

$$q^g_{i,s,t}(h, d) = \mathbb{E}\Big[\overline{Q}^g_{i,s+1,t}(H_{i,s+1}) \mid H_{i,s} = h, D^g_{i,s,t} = d\Big], \tag{24}$$

$$\overline{Q}^g_{i,s,t}(h) = \int q^g_{i,s,t}(h, d) g_{i,s,t}(dd \mid h). \tag{25}$$

Proposition B.8 (Identification of a focal regime value)**.** *Under the conditions above,*

$$\Psi^g_t = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} \overline{Q}^g_{i,s_0,t}(H_{i,s_0}) \middle| \mathcal{D}\right].$$

Thus the four values $\Psi^{00}_t, \Psi^{10}_t, \Psi^{01}_t, \Psi^{11}_t$ and all contrasts in Section 3.2 are identified whenever the conditions hold for all four regimes.

Proof. Fix a focal unit and suppress r, i. Let

$$m_t(h, d) = \mathbb{E}(Y_t \mid H_t = h, D^g_{t,t} = d)$$

and, for $s < t$, let

$$K_s(dh' \mid h, d) = \mathbb{P}(H_{s+1} \in dh' \mid H_s = h, D^g_{s,t} = d)$$

be the observed terminal-outcome and state-transition kernels after the natural complement has been marginalized under $p(c \mid h, d)$. Actionable positivity makes these kernels identified for policy-supported actions at almost every regime-reachable history. To justify this use of observed kernels, the baseline history law is unchanged before s_0. If the intervened history law at s is absolutely continuous with respect to the observed one, assignment positivity preserves absolute continuity after drawing D_s from g_s. Retaining the same natural-complement and state-transition kernels preserves it at $s + 1$. Forward induction thus gives $P^g_{H_s} \ll P_{H_s}$ throughout the finite window. Choices of versions of the observed conditional kernels on observed null sets therefore do not change the regime functional. Conditional identities in the backward induction hold in this almost-everywhere sense.

We prove by backward induction that $\overline{Q}^g_{s,t}(h)$ equals the conditional mean of the terminal potential outcome obtained by following g from time s onward, given that the partial-regime history at s is h. At $s = t$, consistency and sequential exchangeability give

$$\mathbb{E}(Y^d_t \mid H_t = h) = m_t(h, d)$$

for every d in the support of $g_{t,t}(\cdot \mid h)$. Therefore

$$\mathbb{E}(Y^{g_t}_t \mid H^g_t = h) = \int m_t(h, d) g_{t,t}(dd \mid h) = \overline{Q}^g_{t,t}(h),$$

which establishes the base case.

Now suppose the assertion holds at $s + 1$. Consistency and sequential exchangeability identify the conditional law of $H^{\bar{d}_{s_0:s-1}, d_s}_{s+1}$, given $H_s = h$, with $K_s(\cdot \mid h, d_s)$. The induction hypothesis and iterated expectation then yield

$$\begin{aligned}
&\mathbb{E}\Big(Y^{\bar{d}_{s_0:s-1}, g_{s:t}}_t \mid H^{\bar{d}_{s_0:s-1}}_s = h\Big) \\
&\quad = \int\int \overline{Q}^g_{s+1,t}(h') K_s(dh' \mid h, d) g_{s,t}(dd \mid h) \\
&\quad = \int q^g_{s,t}(h, d) g_{s,t}(dd \mid h) = \overline{Q}^g_{s,t}(h).
\end{aligned}$$

Thus the assertion holds down to s_0. Averaging the resulting conditional mean over the observed baseline history and then over focal units gives the displayed functional for Ψ^g_t. Exposure sufficiency only permits the transition and outcome regressions to compress peer history through $E^\star$; the induction itself remains on the actionable-assignment regime and therefore still requires actionable-assignment positivity. The natural complement is integrated into m_t and K_s as in Eq. (23). □

B.3 Longitudinal augmented score

Define the step and cumulative assignment ratios

$$\rho^g_{i,s,t} = \frac{g_{i,s,t}(D^g_{i,s,t} \mid H_{i,s})}{p^D_{i,s,t}(D^g_{i,s,t} \mid H_{i,s})}, \qquad W^g_{i,s,t} = \prod_{u=s_0}^{s} \rho^g_{i,u,t}.$$

The Radon–Nikodym derivative of $g(d \mid h)p(c \mid h, d)$ relative to $p^D(d \mid h)p(c \mid h, d)$ is exactly $g(d \mid h)/p^D(d \mid h)$; the natural-complement factor cancels. To make dependence on a learned exposure map explicit, let T_ϕ be the deterministic transformation that maps a raw history and actionable assignment to the common feature domain used by the fitted nuisance functions. A fitted collection $\nu = (q, p)$ is held fixed as a set of functions on that common domain; $\psi(O; \nu, \phi)$ means that these fixed functions are evaluated after applying T_ϕ. This supplies a common coupling for comparing $\widehat{\phi}$ with $\phi^\star$, even when the fitted functions were trained using $\widehat{\phi}$.

For every candidate outcome collection $q = (q_{s_0}, \ldots, q_t)$, define its regime-integrated companions by

$$\begin{aligned}
\overline{Q}_{t+1}(q) &= Y_t, \\
\overline{Q}_s(q)(h) &= \int q_s(h, d) g_{s,t}(dd \mid h), \qquad s = t, \ldots, s_0.
\end{aligned} \tag{26}$$

We call $(q, \overline{Q}(q))$ *regime compatible*. Every score below uses these integrated companions. We first analyze exact integration; Proposition B.10 treats its Monte Carlo approximation. Define

$$\psi^g_{r,i,t}(O_r; \nu, \phi^\star) = \overline{Q}^g_{i,s_0,t}(H_{i,s_0}) \tag{27}$$

$$+ \sum_{s=s_0}^{t} W^g_{i,s,t} \left\{ \overline{Q}^g_{i,s+1,t}(H_{i,s+1}) - q^g_{i,s,t}(H_{i,s}, D^g_{i,s,t}) \right\}, \tag{28}$$

where the continuation at $s = t$ is $Y_{i,t}$, and let $\psi^g_{r,t} = N^{-1} \sum_i \psi^g_{r,i,t}$.

Proposition B.9 (Mean of the augmented score)**.** *At the true nuisance functions $\nu_0 = (q_0, p_0)$,*

$$\mathbb{E}\{\psi^g_{r,t}(O_r; \nu_0, \phi^\star) \mid \mathcal{D}\} = \Psi^g_t.$$

Proof. For each s, the continuation residual $\overline{Q}^g_{i,s+1,t} - q^g_{i,s,t}$ has conditional mean zero given $(H_{i,s}, D^g_{i,s,t})$. Multiplication by the measurable cumulative ratio preserves this zero conditional mean. Sequential change of measure under $W^g_{i,s,t}$ then telescopes the regime-integrated continuation values back to $\overline{Q}^g_{i,s_0,t}$. Proposition B.8 identifies the expectation of that baseline value with Ψ^g_t. Averaging focal scores within a replicate preserves the identity without node-level independence. □

Proposition B.10 (Finite-M regime-integration error)**.** *Condition on all training folds and held-out observed trajectories. Replace each exact integral in Eq. (26) used for held-out scoring by the average of M conditionally independent draws from $g_{i,s,t}(\cdot \mid H_{i,s})$. Let $\widehat{\Psi}^g_t$ denote the exact-integral cross-fitted estimator and $\widetilde{\Psi}^g_{M,t}$ its Monte Carlo version. Suppose the resulting, within-network averaged score error is conditionally unbiased, has variance at most C/M uniformly over replicates, and its random numbers are independent across network replicates. Then*

$$\widetilde{\Psi}^g_{M,t} - \widehat{\Psi}^g_t = O_p\Big((RM)^{-1/2}\Big). \tag{29}$$

For fixed M, this error is part of the root-R limiting variance; if $M \to \infty$, it is $o_p(R^{-1/2})$. Scoring streams must be independent of the randomization used to fit the nuisances, as well as independent across replicates. Monte Carlo draws used to construct training continuation targets affect the fitted q and are therefore already included in the outcome-nuisance error, not in Eq. (29).

Proof. Let $e_{r,M}$ be the Monte Carlo error of the node-averaged held-out score in replicate r. Conditional on training and held-out data, $\mathbb{E}(e_{r,M}) = 0$, $\mathrm{Var}(e_{r,M}) \le C/M$, and the $e_{r,M}$ are independent across r. Hence

$$\mathrm{Var}\left(R^{-1} \sum_{r=1}^{R} e_{r,M} \right) \le \frac{C}{RM}.$$

Conditional Chebyshev inequality gives Eq. (29); removing the conditioning preserves the stochastic order. Multiplying by $\sqrt{R}$ shows that fixed M leaves a mean-zero term of variance at most C/M, while $M \to \infty$ makes it vanish. □

For focal-array functions, use $\|f\|_2^2 = N^{-1} \sum_i \mathbb{E}[f_i(O)^2]$; the corresponding scalar norm applies when the focal index is suppressed. For a regime-compatible candidate q, define the finite-horizon outcome-nuisance norm

$$\|q - q_0\|^2_{\mathcal{Q},t} := \sum_{s=s_0}^{t} \left(\|q_s - q_{0,s}\|_2^2 + \|\overline{Q}_{s+1} - \overline{Q}_{0,s+1}\|_2^2 \right).$$

For a candidate assignment law p, define the cumulative-ratio norm

$$\|p - p_0\|^2_{\mathcal{W},t} := \sum_{s=s_0}^{t} \|W_s(p) - W_s(p_0)\|_2^2.$$

Under bounded step ratios, step-ratio errors control this cumulative-ratio norm.

Assumption B.11 (Exposure stability)**.** Fold allocation is fixed in advance or randomized independently of the observed trajectories. Let $\mathcal{T}_{-k}$ include the training replicates outside fold k, the fold allocation, and the randomization used to fit that fold's nuisances, and let d_Φ be a metric on exposure maps. For fixed K,

$$\max_{1\le k\le K} d_\Phi(\widehat{\phi}^{(-k)}, \phi^\star) = O_p(\epsilon_\phi), \qquad \epsilon_\phi = o(1).$$

Let $\mathcal{N}(\nu_0)$ be a fixed $L_2(P)$ neighborhood of the true nuisance collection. Require

$$\Pr\Big\{\widehat{\nu}^{(-k)} \in \mathcal{N}(\nu_0) \text{ for every } k = 1, \dots, K\Big\} \longrightarrow 1. \tag{30}$$

With probability tending to one over training and fitting randomization, uniformly in k and over fixed, regime-compatible nuisance functions in $\mathcal{N}(\nu_0)$, the conditionally evaluated replicate score is locally Lipschitz in the exposure map:

$$\big\|\psi^g_{r,t}(O_r; \nu, \phi) - \psi^g_{r,t}(O_r; \nu, \phi^\star)\big\|_{P,2|\mathcal{T}_{-k}} \le L_\phi d_\Phi(\phi, \phi^\star)$$

for a finite constant L_ϕ. Both scores use the common-domain evaluation through T_ϕ defined above. Condition (30) ensures that the Lipschitz bound applies to the actual fitted functions.

Proposition B.12 (Exposure-learning drift)**.** *Under Assumption B.11, replacing $\phi^\star$ by the fold-specific $\widehat{\phi}^{(-k)}$ changes the cross-fitted regime-value estimator by $O_p(\epsilon_\phi)$.*

Proof. Condition on $\mathcal{T}_{-k}$, so the fitted exposure map and nuisance functions are fixed on the held-out replicates. Set

$$D_k(O) = \psi^g(O; \widehat{\nu}^{(-k)}, \widehat{\phi}^{(-k)}) - \psi^g(O; \widehat{\nu}^{(-k)}, \phi^\star).$$

Let $\mathcal{E}_R$ be the event that all fitted nuisances lie in $\mathcal{N}(\nu_0)$ and the stated uniform Lipschitz bounds hold. Assumption B.11 gives $\Pr(\mathcal{E}_R) \to 1$. On each fold's corresponding training-measurable event, the bound applies to $\widehat{\nu}^{(-k)}$; Cauchy–Schwarz gives

$$|PD_k| \le L_\phi d_\Phi(\widehat{\phi}^{(-k)}, \phi^\star) = O_p(\epsilon_\phi).$$

Moreover,

$$\mathrm{Var}\{(P_{R,k} - P)D_k \mid \mathcal{T}_{-k}\} \le R_k^{-1}\|D_k\|^2_{P,2|\mathcal{T}_{-k}} = O_p(\epsilon_\phi^2/R_k).$$

Conditional Chebyshev inequality therefore gives $(P_{R,k} - P)D_k = O_p(\epsilon_\phi R_k^{-1/2})$, and hence $P_{R,k}D_k = O_p(\epsilon_\phi)$. Fixed-$K$, fold-size-weighted aggregation preserves this order. The complement of $\mathcal{E}_R$ has probability $o(1)$, so it does not change these stochastic-order conclusions. □

Theorem B.13 (Cross-fitted orthogonal expansion)**.** *Keep the network size N, target horizon, and number of target regimes fixed as $R \to \infty$. Let $\widehat{q}^{(-k)}$, $\widehat{p}^{(-k)}$, and $\widehat{\phi}^{(-k)}$ be fitted outside replicate fold k, with every candidate outcome collection satisfying Eq. (26). Let $R_k = |\mathcal{I}_k|$, assume K is fixed and $R_k/R \to \pi_k \in (0,1)$, and define*

$$\mathfrak{R}_{q,p,t} := \sum_{k=1}^{K} \frac{R_k}{R}\|\widehat{q}^{(-k)} - q_0\|_{\mathcal{Q},t}\|\widehat{p}^{(-k)} - p_0\|_{\mathcal{W},t}. \tag{31}$$

Suppose Assumptions B.1–B.7 and Assumption B.11 hold. The regime g is fixed independently of p_0 and its estimates, and the theoretical ratios are not truncated. Fitted nodewise probabilities may be projected to $[c, 1-c]$ as in Assumption B.4; cumulative weights are not capped. Finally, suppose the true score has a bounded second moment and

$$\max_{1\le k\le K} \|\psi(\widehat{\nu}^{(-k)}, \phi^\star) - \psi(\nu_0, \phi^\star)\|_{P,2} = o_p(1).$$

Then

$$\begin{aligned}\widehat{\Psi}^g_t - \Psi^g_t = \frac{1}{R}\sum_{r=1}^{R} \varphi^g_t(O_r) + O_p(\mathfrak{R}_{q,p,t}) \\ + O_p(\epsilon_\phi) + o_p(R^{-1/2}),\end{aligned} \tag{32}$$

where $\varphi^g_t(O_r) = \psi^g_{r,t}(O_r; \nu_0, \phi^\star) - \Psi^g_t$. Thus, fold by fold and with $\phi^\star$ fixed, the population bias is exactly zero when either the compatible backward-Q collection or the assignment-law nuisance is correct.

Proof. First hold the exposure map at $\phi^\star$. For fold k, condition on $\mathcal{T}_{-k}$, so $\widehat{\nu}_k = (\widehat{q}^{(-k)}, \widehat{p}^{(-k)})$ is fixed on the held-out observations. Writing $P_{R,k}$ for the empirical measure on fold k and P for the fresh-replicate population measure, add and subtract the true score:

$$\begin{aligned} P_{R,k}\psi(\widehat{\nu}_k, \phi^\star) - \Psi_t^g &= (P_{R,k} - P)\psi(\nu_0, \phi^\star) \\ &\quad + (P_{R,k} - P)\{\psi(\widehat{\nu}_k, \phi^\star) - \psi(\nu_0, \phi^\star)\} \\ &\quad + P\{\psi(\widehat{\nu}_k, \phi^\star) - \psi(\nu_0, \phi^\star)\}. \end{aligned} \tag{33}$$

Cross-fitting makes the nuisance functions fixed when the held-out empirical process is evaluated. Its conditional variance is bounded by

$$R_k^{-1}\|\psi(\widehat{\nu}_k, \phi^\star) - \psi(\nu_0, \phi^\star)\|_2^2.$$

The uniform $L_2(P)$ convergence, $R_k/R \to \pi_k > 0$, and conditional Chebyshev inequality make the second line $o_p(R_k^{-1/2}) = o_p(R^{-1/2})$. Because K is fixed, its fold-size-weighted aggregate is also $o_p(R^{-1/2})$. The weighted aggregate of the first line is exactly

$$\sum_{k=1}^{K} \frac{R_k}{R}(P_{R,k} - P)\psi(\nu_0, \phi^\star) = \frac{1}{R}\sum_{r=1}^{R} \varphi_t^g(O_r).$$

It remains to evaluate the last line fold by fold. Suppress k and the focal-unit index, and write $\Delta q_s = \widehat{q}_s - q_{0,s}$, $\Delta\overline{Q}_s = \overline{Q}_s(\widehat{q}) - \overline{Q}_s(q_0)$, $\widehat{W}_s = W_s(\widehat{p})$, and $W_{0,s} = W_s(p_0)$. By compatibility, $\Delta\overline{Q}_{t+1} = 0$ and

$$\Delta\overline{Q}_s(h) = \int \Delta q_s(h, d) g_{s,t}(dd \mid h). \tag{34}$$

The true continuation residual has conditional mean zero given $(H_s, D_{s,t}^g)$, while $\widehat{W}_s - W_{0,s}$ is measurable with respect to that sigma-field. Therefore

$$P\Big[(\widehat{W}_s - W_{0,s})\{\overline{Q}_{0,s+1} - q_{0,s}\}\Big] = 0.$$

For the candidate errors, one-step change of measure and Eq. (34) give, with $W_{0,s_0-1} \equiv 1$,

$$\begin{aligned} P\{W_{0,s}\Delta q_s(H_s, D_{s,t}^g)\} &= P\big[W_{0,s-1}\mathbb{E}_{p_0}\{\rho_{0,s}\Delta q_s(H_s, D_{s,t}^g) \mid H_s\}\big] \\ &= P\{W_{0,s-1}\Delta\overline{Q}_s(H_s)\}. \end{aligned} \tag{35}$$

Consequently,

$$\begin{aligned} &P\Delta\overline{Q}_{s_0} + \sum_{s=s_0}^{t} P\big[W_{0,s}\{\Delta\overline{Q}_{s+1} - \Delta q_s\}\big] \\ &= P\Delta\overline{Q}_{s_0} + \sum_{s=s_0}^{t} P(W_{0,s}\Delta\overline{Q}_{s+1}) - \sum_{s=s_0}^{t} P(W_{0,s-1}\Delta\overline{Q}_s) = 0, \end{aligned}$$

where every intermediate term cancels and the terminal term is zero because $\Delta\overline{Q}_{t+1} = 0$. Substituting the two preceding identities into the last line of Eq. (33) yields the exact bias representation

$$P\{\psi(\widehat{\nu}, \phi^\star) - \psi(\nu_0, \phi^\star)\} = \sum_{s=s_0}^{t} P\Big[(\widehat{W}_s - W_{0,s})\{\Delta\overline{Q}_{s+1} - \Delta q_s\}\Big]. \tag{36}$$

Cauchy–Schwarz across observations and then across the finite time indices gives

$$\begin{aligned} &|P\{\psi(\widehat{\nu}, \phi^\star) - \psi(\nu_0, \phi^\star)\}| \\ &\quad \le \left\{\sum_{s=s_0}^{t} \|\widehat{W}_s - W_{0,s}\|_2^2\right\}^{1/2} \left\{\sum_{s=s_0}^{t} \|\Delta\overline{Q}_{s+1} - \Delta q_s\|_2^2\right\}^{1/2} \\ &\quad \le \sqrt{2}\,\|\widehat{q} - q_0\|_{\mathcal{Q},t}\|\widehat{p} - p_0\|_{\mathcal{W},t}. \end{aligned}$$

Thus the finite-horizon definitions of $\|\cdot\|_{\mathcal{Q},t}$ and $\|\cdot\|_{\mathcal{W},t}$ yield exactly the fold-specific factor in Eq. (31). Weighting by R_k/R and summing the fixed number of folds gives $O_p(\mathfrak{R}_{q,p,t})$. Proposition B.12 adds $O_p(\epsilon_\phi)$ when the common-domain score is evaluated at $\widehat{\phi}^{(-k)}$ rather than $\phi^\star$. Combining these terms with the empirical-process decomposition proves Eq. (32). If $\widehat{p}^{(-k)} = p_0$, the weight difference in Eq. (36) is zero; if $\widehat{q}^{(-k)} = q_0$, compatibility makes both candidate-error terms zero. This proves the stated fold-wise double-robustness claim. □

Corollary B.14 (Neyman orthogonality at the reference exposure)**.** *Keep g and $\phi^\star$ fixed. Let q_ϵ, p_ϵ be a regime-compatible path through q_0, p_0 such that $\|q_\epsilon - q_0\|_{\mathcal{Q},t} = O(|\epsilon|)$ and $\|p_\epsilon - p_0\|_{\mathcal{W},t} = O(|\epsilon|)$, with the integrability conditions of Theorem B.13. Then*

$$\frac{d}{d\epsilon} P\psi(q_\epsilon, p_\epsilon, \phi^\star)\Big|_{\epsilon=0} = 0.$$

Proof. Equation (36) and its Cauchy–Schwarz bound give $|P\psi(q_\epsilon, p_\epsilon, \phi^\star) - \Psi_t^g| = O(\epsilon^2)$. Dividing by ϵ and taking its two-sided limit proves the derivative statement. A perturbation of only one nuisance has exactly zero population bias when the other stays at truth. □

Proposition B.15 (Cumulative-weight truncation)**.** *For fitted, regime-compatible nuisances, replace $\widehat{W}_s$ in the exact-integral score by $\widehat{W}_{s,b} = \min(\widehat{W}_s, b)$, with $b > 0$. Write $\widehat{\Psi}_{t,b}^g$ for the resulting estimator and define*

$$r_{r,i,s} = \widehat{\overline{Q}}_{r,i,s+1} - \widehat{q}_{r,i,s}, \qquad d_{r,i,s,b} = (\widehat{W}_{r,i,s} - b)_+.$$

Then

$$\widehat{\Psi}_{t,b}^g - \widehat{\Psi}_t^g = -\frac{1}{RN} \sum_{r,i} \sum_{s=s_0}^{t} d_{r,i,s,b} r_{r,i,s}, \tag{37}$$

$$|\widehat{\Psi}_{t,b}^g - \widehat{\Psi}_t^g| \le \mathfrak{C}_{b,t} := \sum_{s=s_0}^{t} \left\{ \frac{1}{RN} \sum_{r,i} d_{r,i,s,b}^2 \right\}^{1/2} \left\{ \frac{1}{RN} \sum_{r,i} r_{r,i,s}^2 \right\}^{1/2}. \tag{38}$$

The untruncated expansion therefore acquires an additional $O_p(\mathfrak{C}_{b,t})$ term. A fixed cap does not in general preserve double robustness with respect to a correct assignment model alone.

Proof. Subtract the two scores; their baseline terms agree and only the cumulative weights differ. Since $\min(w, b) - w = -(w - b)_+$, this gives Eq. (37). The triangle inequality followed by Cauchy–Schwarz over (r, i) gives Eq. (38). When the continuation regression is correct, its residual has conditional mean zero given the full history and current action, so multiplying by a measurable capped weight preserves this zero mean. When only the assignment law is correct, the capped weights no longer implement the change of measure used by the telescoping identity. The resulting bias need not vanish. For a binary action with $p(D = 1) = 1/4$, $g(D = 1) = 3/4$, $Y = D$, $q \equiv 0$, and $b = 1$, the target is $3/4$ but the capped score has mean $1/4$. Both regimes have observed support. □

Corollary B.16 (Randomized finite-M expansion)**.** *Under Theorem B.13 and Proposition B.10,*

$$\widetilde{\Psi}_{M,t}^g - \Psi_t^g = \frac{1}{R} \sum_{r=1}^{R} \varphi_t^g(O_r) + O_p(\mathfrak{R}_{q,p,t}) + O_p(\epsilon_\phi) + O_p((RM)^{-1/2}) + o_p(R^{-1/2}). \tag{39}$$

If $M \to \infty$, the exact-integral limit theory is unchanged. If M is fixed, a central limit theorem additionally requires the convergence and remainder-rate conditions of Theorem B.18. Its variance includes the Monte Carlo score error. For capped scores, Proposition B.15 also requires negligible truncation error before this limit can be used.

Corollary B.17 (Effect contrasts and replicate-level limit)**.** *Apply Theorem B.13 jointly to $g^{00}, g^{10}, g^{01}, g^{11}$. Any fixed linear contrast $\Theta_t = c^\top(\Psi_t^{00}, \Psi_t^{10}, \Psi_t^{01}, \Psi_t^{11})$, including IDE, IPE, ITE, and INT, has the expansion obtained by applying $c^\top$ to the four regime-specific influence functions and remainders. If every fold-wise product remainder and every exposure-drift term is $o_p(R^{-1/2})$, and the resulting replicate influence function has finite, nonzero variance $\sigma_{c,t}^2$, then*

$$\sqrt{R}(\widehat{\Theta}_t - \Theta_t) \rightsquigarrow N(0, \sigma_{c,t}^2) \qquad (R \to \infty).$$

For any fixed finite collection of target times, the corresponding trajectory vector obeys the analogous multivariate limit whenever its replicate influence-vector covariance matrix is finite.

Proof. Linearity of the four regime-value estimators gives the stated contrast expansion. Under the displayed rate conditions, all non-influence terms are $o_p(R^{-1/2})$. Assumption B.7 and the ordinary i.i.d. central limit theorem apply to the centered replicate-level contrast influence function; Slutsky's theorem gives the univariate limit. The multivariate statement follows by the Cramér–Wold device. □

Theorem B.18 (Joint fixed-M limit and contrast variance)**.** *Assume all conditions of Theorem B.13 and Proposition B.10 hold jointly for the four regimes below, with common replicate folds and no cumulative-weight cap. In particular, retain the exact-score convergence requirement*

$$\max_{g \in \{g^{00}, g^{10}, g^{01}, g^{11}\}} \max_{1 \le k \le K} \|\psi_t^g(\widehat{\nu}^{g,(-k)}, \phi^\star) - \psi_t^g(\nu_0^g, \phi^\star)\|_{P,2} = o_p(1).$$

Here ν^g includes the regime-specific continuation functions and the relevant assignment law. Order the four regimes as $(g^{00}, g^{10}, g^{01}, g^{11})$ and write $\mathbf{\Psi}_t$, $\widehat{\mathbf{\Psi}}_t$, and $\widetilde{\mathbf{\Psi}}_{M,t}$ for the corresponding value vectors. Let U_r collect the base random numbers used for held-out regime integration in replicate r. These random-number arrays are identically distributed and independent across replicates, independent of all observed trajectories, and independent of fold allocation and nuisance-training randomization. Base random numbers may be shared across regimes within a replicate. Actual assignment draws are measurable transformations

$$D_{r,i,s}^{g,(m)} = G_{i,s,t}^g(H_{r,i,s}, U_r^{(m)}), \qquad D_{r,i,s}^{g,(m)} \mid H_{r,i,s} \sim g_{i,s,t}(\cdot \mid H_{r,i,s}),$$

where $U_r^{(m)}$ denotes the relevant random-number subarray and G^g is a sampler for the specified regime. At the true nuisances, define the four-vector Monte Carlo score error

$$\boldsymbol{e}_{r,M} := \widetilde{\boldsymbol{\psi}}_{t,M}(O_r, U_r; \nu_0, \phi^\star) - \boldsymbol{\psi}_t(O_r; \nu_0, \phi^\star).$$

Suppose, for some $\zeta > 0$,

$$\mathbb{E}(\boldsymbol{e}_{r,M} \mid O_r) = \mathbf{0}, \qquad \mathbb{E}\|\boldsymbol{e}_{r,M}\|^{2+\zeta} < \infty, \tag{40}$$

and, writing $\widehat{\boldsymbol{e}}_{k,M}$ for the randomized fitted score minus its exact-integral counterpart in fold k, require

$$\max_{k \le K} \|\widehat{\boldsymbol{e}}_{k,M} - \boldsymbol{e}_M\|_{L_2(P \otimes P_U | \mathcal{T}_{-k})} = o_p(1).$$

Assume every product and exposure-drift remainder is $o_p(R^{-1/2})$ and the exact influence vector $\boldsymbol{\varphi}_t(O_r)$ has a finite $(2+\zeta)$ moment. Then, for fixed M,

$$\sqrt{R}\left(\widetilde{\mathbf{\Psi}}_{M,t} - \mathbf{\Psi}_t\right) \rightsquigarrow N_4(\mathbf{0}, \mathbf{\Sigma}_t + \mathbf{\Omega}_{M,t}), \tag{41}$$

where $\mathbf{\Sigma}_t = \mathrm{Var}\{\boldsymbol{\varphi}_t(O_r)\}$ and $\mathbf{\Omega}_{M,t} = \mathbb{E}(\boldsymbol{e}_{r,M}\boldsymbol{e}_{r,M}^\top)$. With M independent draws per regime integral, $\mathbf{\Omega}_{M,t} = O(M^{-1})$; off-diagonal entries retain cross-regime common-random-number covariance.

For any fixed contrast vector c, including IDE, IPE, ITE, and INT,

$$\sqrt{R}\left\{c^\top \widetilde{\mathbf{\Psi}}_{M,t} - c^\top \mathbf{\Psi}_t\right\} \rightsquigarrow N\big(0, c^\top(\mathbf{\Sigma}_t + \mathbf{\Omega}_{M,t})c\big). \tag{42}$$

Let $\widetilde{\boldsymbol{\psi}}_{r,M,t}$ be the four held-out, node-averaged implemented scores and let $\widetilde{\mathbf{\Psi}}_{M,t} = R^{-1}\sum_r \widetilde{\boldsymbol{\psi}}_{r,M,t}$. A consistent variance estimator for the contrast estimate is

$$\widehat{\mathrm{Var}}\Big(c^\top \widetilde{\mathbf{\Psi}}_{M,t}\Big) = \frac{1}{R(R-1)} \sum_{r=1}^{R} \left[c^\top \left\{\widetilde{\boldsymbol{\psi}}_{r,M,t} - \widetilde{\mathbf{\Psi}}_{M,t}\right\}\right]^2. \tag{43}$$

It includes sampling and Monte Carlo variation and all four-regime covariances.

Proof. The inherited exact-score convergence and cross-fitting conditions justify applying Corollary B.16 componentwise. By the assumed L_2 convergence of the randomized fitted score, replicate-level cross-fitting and conditional Chebyshev inequality replace its Monte Carlo error by $\boldsymbol{e}_{r,M}$ with an aggregate $o_p(R^{-1/2})$ difference. Specifically, conditional unbiasedness of both integration errors implies $\mathbb{E}_U[\widehat{\boldsymbol{e}}_{k,M} - \boldsymbol{e}_M \mid O, \mathcal{T}_{-k}] = 0$; their conditional L_2 difference bounds the variance of the held-out fold average by $o_p(1)/R_k$. Fixed-K aggregation therefore has the asserted order, without requiring independence between training folds. The stated remainder rates give

$$\sqrt{R}\left(\widetilde{\boldsymbol{\Psi}}_{M,t} - \boldsymbol{\Psi}_t\right) = R^{-1/2}\sum_{r=1}^{R}\{\boldsymbol{\varphi}_t(O_r) + \boldsymbol{e}_{r,M}\} + o_p(1).$$

The pairs (O_r, U_r) are independent and identically distributed across network replicates. The $(2+\zeta)$ moments imply the multivariate Lindeberg condition, so the multivariate central limit theorem applies. Moreover,

$$\mathbb{E}\big[\boldsymbol{\varphi}_t(O_r)\boldsymbol{e}_{r,M}^\top\big] = \mathbb{E}\big[\boldsymbol{\varphi}_t(O_r)\mathbb{E}(\boldsymbol{e}_{r,M}^\top \mid O_r)\big] = 0.$$

Hence the covariance is $\boldsymbol{\Sigma}_t + \boldsymbol{\Omega}_{M,t}$, including off-diagonal Monte Carlo covariance. Cramér–Wold gives Eq. (42). Finally, L_2 convergence of the cross-fitted implemented scores and the law of large numbers make their sample covariance converge to $\boldsymbol{\Sigma}_t + \boldsymbol{\Omega}_{M,t}$. Applying the quadratic form $c^\top(\cdot)c/R$ yields Eq. (43). □

B.4 Replicate-level cross-fitting and inference

Partition the R network trajectories into folds $\mathcal{I}_1, \ldots, \mathcal{I}_K$. For each k, train the DyNICL exposure/representation nuisance, assignment nuisance, and neural backward- Q heads only on $\mathcal{I}_k^c$. Freeze all fitted objects, compute (28) on $r \in \mathcal{I}_k$, first average nodes within each network, and then form

$$\widehat{\Psi}_t^g = \frac{1}{R}\sum_{k=1}^{K}\sum_{r\in\mathcal{I}_k}\psi_{r,t}^g\Big(O_r; \widehat{\nu}^{(-k)}, \widehat{\phi}^{(-k)}\Big).$$

This display defines the exact-regime-integral augmented estimator. Its Monte Carlo version $\widetilde{\Psi}_{M,t}^g$ replaces the held-out regime integrals by M draws and satisfies Proposition B.10 and Corollary B.16 when their randomization conditions hold. Corollary B.17 gives the exact-integral limit, while Theorem B.18 gives the joint four-regime limit and contrast variance for the implemented fixed-M estimator as $R \to \infty$. The independent units in these limits are complete network trajectories. Appendix B.6 gives the Q-based guarantee.

B.5 Deterministic-regime boundary

All-treated and all-control policies are degenerate members of the regime family. They are observationally identified only when each required actionable assignment has positive conditional probability at almost every history under the corresponding regime-induced law.

B.6 Consistency of Q-based regime-value estimation

Fix a target time and regime. Work under $P_{\mathcal{D}}$ as in Appendix B, and let the R trajectory replicates be independent with common law P. Partition their indices independently of their values into a fixed number K of folds I_k, with $R_k/R \to \pi_k > 0$. Write $\pi_{k,R} = R_k/R$ and let $P_{R,k}$ be the held-out fold mean. Define

$$m_0(O) = N^{-1}\sum_i \overline{Q}_{0,i,s_0,t}^g(H_{i,s_0}), \qquad \widehat{m}_k(O) = N^{-1}\sum_i \widehat{\overline{Q}}_{i,s_0,t}^{g,(-k)}(H_{i,s_0}).$$

Training randomization is independent of held-out trajectories, and $\widehat{m}_k$ is measurable with respect to the training fold and that randomization.

Proposition B.19 (Q-based error decomposition)**.** *Suppose* $Pm_0 = \Psi_t^g$, $Pm_0^2 < \infty$, *and* $a_R = \max_k \|\widehat{m}_k - m_0\|_{P,2|\mathcal{T}_{-k}} = o_p(1)$, *where the conditioning includes training randomization. Then*

$$\widehat{\Psi}_t^{g,\mathrm{Q}} - \Psi_t^g = (P_R - P)m_0 + \sum_{k=1}^{K}\pi_{k,R}P(\widehat{m}_k - m_0) + o_p(R^{-1/2}), \tag{44}$$

and the estimator is consistent.

Proof. Adding and subtracting m_0 within every fold gives the exact identity

$$\widehat{\Psi}_t^{g,\mathrm{Q}} - \Psi_t^g = (P_R - P)m_0 + \sum_k \pi_{k,R} P(\widehat{m}_k - m_0) + \sum_k \pi_{k,R}(P_{R,k} - P)(\widehat{m}_k - m_0).$$

Conditional on the training data for fold k, the last fold-specific term before multiplication by $\pi_{k,R}$ has mean zero and variance at most $\|\widehat{m}_k - m_0\|^2_{P,2|\mathcal{T}_{-k}}/R_k$. For any $\epsilon, \delta > 0$, conditional Chebyshev and a split according to whether this norm exceeds δ give

$$\Pr\Big(\sqrt{R}\,|(P_{R,k} - P)(\widehat{m}_k - m_0)| > \epsilon\Big) \le \Pr\big(\|\widehat{m}_k - m_0\|_{P,2|\mathcal{T}_{-k}} > \delta\big) + \frac{R\delta^2}{R_k\epsilon^2}.$$

Taking the large-R limit and then $\delta \downarrow 0$ proves that each term is $o_p(R^{-1/2})$. A finite sum has the same order; no independence between the fitted models for different folds is needed. Furthermore, conditional Cauchy–Schwarz bounds the absolute mean-error term by $\sum_k \pi_{k,R}\|\widehat{m}_k - m_0\|_{P,2|\mathcal{T}_{-k}} \le a_R$. The first term converges to zero by the law of large numbers, and the other two terms converge to zero by the preceding bounds. This proves both the decomposition and consistency. □

The generally first-order mean-continuation term in Eq. (44) requires control for root-R inference. Convergence of $\widehat{m}_k$ covers exposure, representation, regression, and training-time integration errors.

C Detailed Model Implementation

C.1 Exposure Features and Numerical Conventions

The fold-specific learned map in Section 4 is

$$\widehat{E}^{(-k)}_{r,i,s} = \widehat{\phi}^{(-k)}_i(\mathbf{A}_{\mathrm{obs}}, \mathbf{Z}_{r,S_i,s-1}, \ldots, \mathbf{Z}_{r,S_i,s-\tau_{\max}}). \tag{45}$$

For tensor consistency the implementation retains an identically zero hop-0 block, $B^{(0,\tau)} = \widehat{E}^{(0,\tau)} = 0$ and $\alpha_{0,\tau} = 0$. Its feature vector is

$$\widehat{\mathbf{E}}^{\mathrm{feat}}_{i,s} = \mathrm{Concat}_{h=0:H,\,\tau=1:\tau_{\max}} \widehat{E}^{(h,\tau)}_{i,s}.$$

The token summary averages these entries, including the zero block:

$$\overline{E}^{\mathrm{atom}}_{i,s} = \frac{1}{(H+1)\tau_{\max}} \sum_{h=0}^{H} \sum_{\tau=1}^{\tau_{\max}} \widehat{E}^{(h,\tau)}_{i,s}.$$

The scalar mixture in Eq. (12) instead weights raw positive-hop atoms. Numerically it obtains these from the lag-weighted atoms by division by $\max(k_\tau, 10^{-12})$. This is equal to the raw atom when the denominator floor is inactive. Pre-time-zero histories are zero-padded.

Empty rows of W_θ and graph-attention rows without allowed neighbors have zero context. Empty graph support contributes zero logit regularization. The default uses asymmetric edge logits, all allowed neighbors, and dense hop–lag softmax weights. On a binary skeleton, adding a constant to all allowed edges cancels in the row softmax, but penalizing the logits can affect optimization.

C.2 Encoders, Tokens, and Auxiliary Training

Write the focal history and the corresponding peer-history collection as

$$\overline{\mathcal{H}}_{i,s} = \{\mathbf{X}_{i,0:s}, Z_{i,0:s-1}, Y_{i,0:s-1}\}. \tag{46}$$

$$\overline{\mathcal{H}}^{\mathrm{peer}}_{i,s} = \{\overline{\mathcal{H}}_{j,s} : j \in S_i\}.$$

Here S_i is the interference set defined in Section 3.1. The shared encoding receives both histories together with exposure features:

$$(\mathbf{h}_{i,s}^{\mathrm{hist}}, \mathbf{h}_{i,s}^{\mathrm{peer}}) = \mathrm{Enc}_\eta(\overline{\mathcal{H}}_{i,s}, \overline{\mathcal{H}}_{i,s}^{\mathrm{peer}}, \widehat{\mathbf{E}}_{i,s}^{\mathrm{feat}}; W_\theta). \tag{47}$$

Both history inputs precede the current treatment and outcome. Masked dot-product graph attention uses

$$\mathrm{Attn}(Q, K, V; W_\theta) = \mathrm{Softmax}\Big(QK^\top/\sqrt{d} + B(W_\theta) + M(W_\theta)\Big)\, V, \tag{48}$$

with $B_{ij} = \log(W_\theta[i,j] + \epsilon)$ on allowed off-diagonal entries and $M_{ij} = -\infty$ on excluded entries. A residual projection retains local exposure atoms. The three tokens are

$$\begin{aligned} &\mathbf{b}_{i,s} = \varphi_b(\mathbf{h}_{i,s}^{\mathrm{hist}}), \\ &\mathbf{u}_{i,s} = \varphi_u([\mathbf{h}_{i,s}^{\mathrm{hist}}, Z_{i,s-1}]), \qquad\qquad \mathbf{v}_{i,s} = \varphi_v([\mathbf{h}_{i,s}^{\mathrm{peer}}, \overline{E}_{i,s}^{\mathrm{atom}}]). \end{aligned} \tag{49}$$

The mixer blocks direct u–v attention in both directions and allows each token to attend to b. Its three outputs form Eq. (13) together with scalar exposure. Diagonal masking removes each local atom's own-treatment input, but graph attention can return own history through neighbors' atoms.

The auxiliary outcome head is

$$\widehat{Y}_{i,s} = m_\omega(\mathbf{R}_{i,s}^{\mathrm{dir}}, \mathbf{R}_{i,s}^{\mathrm{peer}}, Z_{i,s}) + c_E \widehat{E}_{i,s} + c_{ZE} Z_{i,s} \widehat{E}_{i,s}. \tag{50}$$

Current own treatment enters after encoding; scalar exposure and its interaction supply gradients to the exposure mixture. The balancing penalty is the linear-kernel mean discrepancy

$$\mathcal{L}_{\mathrm{bal}} = \frac{1}{d_R} \left\| \mathrm{Mean}_{Z=1}\, \mathbf{R}^{\mathrm{base}} - \mathrm{Mean}_{Z=0}\, \mathbf{R}^{\mathrm{base}} \right\|_2^2, \tag{51}$$

computed over node–time observations within a training replicate and set to zero when either treatment group is empty. For the allowed-edge set $\mathcal{S}$, exposure regularization is

$$\mathcal{R}_{\mathrm{LEM}} = \lambda_W\, \mathrm{Mean}_{(i,j)\in\mathcal{S}}\, |L_{ij}| - \lambda_\alpha \sum_{h,\tau} \alpha_{h,\tau} \log(\alpha_{h,\tau} + \epsilon). \tag{52}$$

Together with prediction error, these terms give Eq. (14).

C.3 Continuation Fitting and Numerical Evaluation

Each neural head pools focal-node rows within the training replicates for its (s, t, g). The action vector contains ordered actionable peer assignments before t and only own assignment at t, with inactive slots zero-padded. Training-set normalization is frozen for held-out evaluation. Regressing on pre-assignment features and actionable treatment marginalizes the contemporaneous natural complement as in Eq. (7).

For every candidate q_s, its companion $\overline{Q}_s$ is obtained by integration under the same fixed policy.

Exact evaluation enumerates actionable assignments and sums their fitted continuations with policy masses. Monte Carlo evaluation approximates the same integral. Randomized training targets affect the fitted-Q error; randomized held-out integration contributes a separate term only under Proposition B.10. Scoring streams must be independent of fitting randomization and independent across replicates, although a realized policy draw can depend on the observed history.

Projection onto an interval containing the true probability does not increase pointwise probability error. A fixed cumulative cap can add bias even with a correct assignment nuisance (Proposition B.15). The support-sensitivity projection and generating floors are specified in Appendix D.

C.4 Formal configuration and estimator interfaces

The following configuration is held fixed except for the field varied in each sensitivity experiment.

Shared core. The default uses three positive hops, asymmetric edge logits, all allowed proxy neighbors, and dense softmax hop–lag mixing, with the numerical conventions in Appendix C. The shared hidden dimension is 16, with a 60-epoch budget and learning rate 0.003. The balancing coefficient is 0.1; graph-logit and mixing entropy coefficients are both 10^{-4}. The prediction head includes the scalar exposure and own-treatment interaction branch in Eq. (50).

Continuation and assignment nuisances. Each backward head is a two-hidden-layer MLP of width 32 with standardized inputs and targets, an 80-epoch budget, and learning rate 0.003. Its inputs follow Eq. (15), with the joint history interface in Eq. (47) and frozen training-set normalization. The assignment model standardizes the feature vector

$$\left[X_{j,s}, Y_{j,s-1}, Z_{j,s-1}, \widehat{E}_{j,s}, \widehat{E}_{j,s-1}, d_j/d_{\max}, s/(T-1)\right]$$

and fits an ℓ_2-regularized logistic regression with inverse regularization parameter $C = 1$. Missing pre-time-zero entries are zero; degree normalization uses a denominator of at least one. Probabilities are projected to $[10^{-3}, 1 - 10^{-3}]$ by default. The support-sensitivity experiment instead fixes this interval to $[0.01, 0.99]$.

Policies, folds, and integration. The default four policies use own probabilities 0.2 and 0.8 and peer probabilities 0.3 and 0.7. Up to three non-self proxy neighbors are selected in increasing node-index order, before fitting or evaluation. The policy index set is identical across all four cells and methods. For proxy-perturbation comparisons, selection uses the unperturbed design: the actionable set is frozen before adding or deleting estimator-input edges. Three folds split 30 independent network trajectories. Exact integration enumerates at most 2^3 peer assignments before the target and two own assignments at the target, with their policy masses. Cumulative weights are uncapped by default.

Windows and reported targets. The base synthetic design has $N = 64$, $T = 12$; the grid varies degree and reference lag. Its default policy window L_g and model window $\tau_{\max}$ match the declared reference lag δ. Window sensitivity holds the intervention fixed when varying model memory. The estimator computes both full-network means and focal values. For the default focal-trajectory error metric, the truth generator selects 32 focal indices using its recorded random seed, independently of fitted errors. Predictions and oracle values use these same indices. Focal-by-time RMSE is defined in Appendix D.6. Oracle simulation uses 256 draws per regime.

D Experimental Protocol

D.1 Design and causal targets

The main grid crosses expected degree $d \in \{4, 8, 12\}$ with generating memory $L \in \{1, 3, 5\}$. Generating propagation spans one to three hops. Target times are $\{\max(3, L), 7, 11\}$, so the full policy window is available at every evaluated target. The model-range experiment fixes the 8×3 design and intervention window and varies only the learner's maximum hop and lag.

All comparisons use the focal policy contrasts of Section 3.2. Each changed generating mechanism has its own counterfactual truth; input-graph perturbations preserve the original data and target.

D.2 Sequential synthetic data-generating process

Each outer run draws a design $\mathcal{D}$ and then generates independent trajectories conditional on it. The design fixes the reference and proxy graphs, exogenous covariate paths, propagation matrices, hop–lag weights, structural coefficients, and evaluation sets. In the latent-variable diagnostic it also fixes the node effects $\mathbf{U}$. Only treatment randomness and outcome innovations vary across the inner replicates. Oracle values average over these same trajectory innovations, holding $\mathcal{D}$ fixed.

Network design and covariates. An undirected reference adjacency $\mathbf{A}^\star$ is sampled from independent Bernoulli upper-triangular entries and symmetrized. An isolated node is connected to its cyclic successor. The observed proxy is formed by deleting reference edges and adding candidate edges, with zero diagonal. Reference propagation is row-normalized:

$$W^\star[i,j] = \frac{A^\star[i,j]\exp(\ell_{ij})}{\sum_m A^\star[i,m]\exp(\ell_{im})}.$$

The edge logits are bounded Gaussian draws. Reference adjacency is undirected, but weighted propagation need not be symmetric. Node covariates are Gaussian baseline vectors plus a shared sinusoidal time component, $X_{i,s} = x_i + c_X \sin(2\pi s/T)\mathbf{1}$. The assignment and outcome covariate coefficients are drawn once per design.

Exposure from past assignments. Let $H_\star$ be the generating hop limit and δ the generating memory. Normalized geometric hop and lag weights are $v_h \propto r_h^{h-1}$ and $w_\ell \propto r_\ell^{\ell-1}$. For a propagation matrix $W_s^\star$, define

$$E^\star_{r,i,s} = \sum_{h=1}^{H_\star} \sum_{\ell=1}^{\delta} v_h w_\ell \left[P_{h,s}\mathbf{Z}_{r,s-\ell}\right]_i, \qquad \mathbf{Z}_{r,u} = \mathbf{0} \quad (u < 0). \tag{53}$$

The formal experiments use the peer-only operator $P_{h,s} = (W_s^\star)^h \odot (\mathbf{1}\mathbf{1}^\top - I)$; the mask removes return paths from a unit's own past treatment after taking the matrix power, without subsequent row renormalization. In experiments, $W_s^\star = W^\star$ is static and exposure varies with assignment history.

Assignment and outcome equations. Initialize $Y_{r,i,-1} = 0$. At each time s, first compute Eq. (53), then sample treatment independently across nodes conditional on the generating history and design:

$$\begin{aligned} a_{r,i,s} &= a_0 + a_X^\top X_{i,s} + a_E E^\star_{r,i,s} + a_Y Y_{r,i,s-1} + \gamma_Z U_i, \\ p_{r,i,s} &= \epsilon + (1-2\epsilon)\,\mathrm{sigmoid}(a_{r,i,s}), \qquad Z_{r,i,s} \sim \mathrm{Bernoulli}(p_{r,i,s}). \end{aligned} \tag{54}$$

Here U_i is a design-level standard Gaussian variable withheld from the estimator in the latent-variable setting; ϵ sets the generating probability floor. Write $\breve{E}_{r,i,s} = E^\star_{r,i,s} - c_E$. The outcome is

$$\begin{aligned} Y_{r,i,s} = m_s(X_{i,s}, Y_{r,i,s-1}) &+ \beta_Z(s) Z_{r,i,s} + \beta_E(s)\breve{E}_{r,i,s} \\ &+ \beta_{ZE}(s) Z_{r,i,s}\breve{E}_{r,i,s} + \gamma_Y U_i + \lambda_{\mathrm{alt}} E^{\mathrm{alt}}_{r,i,s} + \sigma_Y \xi_{r,i,s}, \end{aligned} \tag{55}$$

where $\xi_{r,i,s}$ are independent standard Gaussian innovations, independent of treatment randomization. The baseline function is

$$m_s(x,y) = b_X^\top x + c_s \sin(2\pi s/T) + b_Y y + \kappa \left\{ \sin(x_1) + c_y \frac{y^2}{1+|y|} \right\}.$$

The three effect coefficients consist of a constant and a periodic component: sine for $\beta_Z(s)$ and $\beta_{ZE}(s)$, and cosine for $\beta_E(s)$. Zero amplitudes give constant coefficients; $\kappa = 0$ removes the nonlinear baseline. The generated outcome $Y_{r,i,s-1}$ drives feedback in both equations.

Boundary settings. The latent-variable diagnostic activates the same unobserved node effect U_i in assignment and outcome. With U_i fixed within each design and withheld from the learner, this tests sensitivity to omitted node heterogeneity. The alternative-exposure term is explicitly $E^{\mathrm{alt}}_{r,i,s} = [W_{\mathrm{alt}}\mathbf{Z}_{r,s-1}]_i$ for $s > 0$, and zero at $s = 0$, where W_{alt} is the row-normalized adjacency of a random relabeling of the reference graph. This term tests omission only to the extent that the estimator's available information cannot represent the alternative exposure. Changing these mechanisms requires recomputing the corresponding causal truth. By contrast, estimator-input graph perturbations preserve the generated trajectories and policy targets as specified below.

Mechanisms used in the reported experiments. The main grid uses three-hop, peer-only propagation with heterogeneous edge weights, geometric hop–lag mixing, a nonlinear outcome baseline, and feedback in both assignment and outcome equations. The peer-effect coefficient varies periodically; the own and interaction coefficients are constant. The exposure is centered before entering the outcome and interaction terms. Latent-variable and alternative-exposure terms are zero in the main, ablation, and hyperparameter experiments. The main proxy has added candidate edges and no reference-edge deletion. The neighborhood/lag grid changes expected degree and generating memory; model-range sensitivity changes only the learner's hop and lag limits. Boundary experiments activate their designated omitted term or change the generating probability floor. Semi-synthetic experiments replace the random graph by empirical contact weights, retaining peer-only static propagation; the short setting disables feedback and the long setting enables it. They contain no latent-variable or alternative-exposure term.

Counterfactual truth. For each focal unit i, target t, and fixed regime g_i^{ab}, the oracle simulates a separate network world from time zero to t. At each step it recomputes exposure, natural assignment

probabilities, and outcomes from that world's history. During the policy window it replaces only the actionable peer draws, and at the target time only the focal treatment draw, by draws from the specified fixed policy. The remaining assignments follow the natural law; its nodewise factorization permits their natural draws to be retained when the actionable components are replaced. For M_{truth} independent rollouts,

$$\widehat{V}_{i,t}^{ab,\mathrm{oracle}} = \frac{1}{M_{\mathrm{truth}}} \sum_{m=1}^{M_{\mathrm{truth}}} Y_{i,t}^{g_i^{ab},(m)}.$$

The four policy worlds use paired random-number streams; their marginal laws remain the specified regimes. Contrasts are computed from these focal means before the metric aggregates focal units and target times. M_{truth} controls oracle approximation and is distinct from the draw count used to integrate a fitted backward-Q function.

D.3 Exposure baselines and module ablations

Exposure baselines share the backward-Q estimation framework and vary the exposure construction. Fixed one-hop uses normalized proxy adjacency with lag one. Fixed multi-hop aggregates one-to-three-hop propagation over the setting's lag window. ObsGraph fixes propagation weights to the proxy structure while learning the remaining exposure aggregation. The oracle-graph reference instead supplies the generating graph.

The module ablations remove learned components of DyNICL. Removing the peer pathway deletes LEM, the downstream peer representation, and scalar exposure route. Removing temporal attention retains the input projection and lagged peer exposure. Removing peer attention retains LEM but deletes its subsequent cross-node aggregation. Removing token mixing retains the three channels but deletes attention-based fusion. All variants use both estimators and the original policy targets.

D.4 Temporal adaptations of external baselines

NetEst (Jiang & Sun, 2022) and TNet (Chen et al., 2024) retain their native outcome heads. Lagged inputs and forward policy rollouts adapt these models to the longitudinal four-cell targets. Training uses the same observed data and replicate folds, and evaluation uses the same focal units, target times, policy probabilities, and truth. Forward policy integration uses 64 Monte Carlo rollouts and natural assignment probabilities clipped to $[0.001, 0.999]$.

NetEst uses hidden width 32, 30 epochs, and learning rate 0.003; TNet uses width 16, 60 epochs, and learning rate 0.0003. Both inherit coefficient settings $\alpha = 0.5$ and $\gamma = 1$ from their development configuration. TNet's targeted update uses learning rate 0.0001, five warm-up epochs, weight decay 0.001, gradient-norm cap 5, and joint-factor floor 0.001. Its targeted optimizer updates only the fluctuation spline. These development configurations are fixed across the reported grid.

D.5 Semi-synthetic contact-network protocol

The SocioPatterns 2012 high-school network has 180 nodes (Fournet & Barrat, 2014). Observed contact topology and frequencies define the network input, while the sequential SCM generates treatment and outcome trajectories over 12 simulated time steps. Each run contains 30 independent trajectories with three replicate folds; 32 pre-selected focal units are evaluated at times $5, \ldots, 11$. The oracle uses 256 paired rollouts per policy, with own probabilities $0.2/0.8$, peer probabilities $0.3/0.7$, and at most three actionable peers. The short-memory case uses lag and policy window one and zero assignment/outcome feedback. The long-memory case uses lag and policy window five, assignment feedback 0.6, and outcome feedback to 0.35. We compare both DyNICL and fixed multi-hop schemes, NetEst, and TNet, without case-specific retuning.

D.6 METRICS AND REPEATED EXPERIMENTS

Let $\theta_{i,t}$ be one of the IDE, IPE, ITE, or INT contrasts of the focal values in Eq. (3). For an evaluated focal set $\mathcal{I}$ and target-time set $\mathcal{T}$, run b yields

$$\xi_b = \left[\frac{1}{|\mathcal{I}||\mathcal{T}|} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} (\widehat{\theta}_{b,i,t} - \theta_{b,i,t})^2 \right]^{1/2}. \tag{56}$$

For $B = 10$ independent repetitions, the reported mean and standard deviation are

$$\overline{\xi} = B^{-1} \sum_{b=1}^{B} \xi_b, \qquad s_\xi = \left[\frac{1}{B-1} \sum_{b=1}^{B} (\xi_b - \overline{\xi})^2 \right]^{1/2}. \tag{57}$$

Here s_ξ measures between-run variation of focal-by-time RMSE. The R network trajectories within each run are the cross-fitting units; the B independent runs determine the reported dispersion.

D.7 SENSITIVITY AND PERTURBATION PROTOCOLS

The hop–lag sensitivity grid varies model maxima over $\{1, 3, 5\}^2$ on fixed 8×3 data. Regularization sensitivities multiply the frozen default balancing or LEM regularization coefficient by 0, 0.1, 1, 10, or 100, holding the other settings fixed.

Proxy perturbations delete or add edges separately at rates 0, 0.1, 0.3, and 0.5. Both rates use the original graph's edge count as denominator. Only the graph supplied to the estimator changes. The generating graph, observations, intervention target, and truth remain fixed, and DyNICL and fixed multi-hop receive the same perturbed graph. Actionable peer sets, policy probabilities, and evaluated focal units are fixed from the unperturbed design before graph perturbation.

Assignment support is varied on 8×3 and 12×5 with generating probability floors 0.20, 0.10, 0.03, and 0.005. Estimated propensities are clipped to $[0.01, 0.99]$ throughout; clipping can introduce bias when the true probability falls outside this interval.

The unobserved-node-effect diagnostic uses strengths 0, 0.25, 0.5, and 1 for the common assignment/outcome term U_i in Eqs. (54)–(55). The omitted-exposure diagnostic varies λ_{alt} over the same grid. Both withhold the added component from the fitted models; the design-conditional interpretation is specified in Appendix D.2.

E COMPLETE SYNTHETIC RESULTS

Q and Aug denote Q-based and augmented estimation. The aggregation convention is defined in Section 5. External baselines use the adaptations in Section D.4.

E.1 NEIGHBOR-COUNT AND LAG GRID

The oracle-graph rows supply the generating graph as additional information.

Table 4: Effect estimation with 4 neighbors and lag 1.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.2547_{\pm 0.0473}$	$0.2638_{\pm 0.0446}$	$0.4016_{\pm 0.0637}$	$0.1964_{\pm 0.0358}$
Fixed one-hop	Aug	$0.2461_{\pm 0.0418}$	$0.2473_{\pm 0.0412}$	$0.3792_{\pm 0.0564}$	$0.1897_{\pm 0.0341}$
Fixed multi-hop	Q	$0.2379_{\pm 0.0436}$	$0.2284_{\pm 0.0397}$	$0.3631_{\pm 0.0568}$	$0.1726_{\pm 0.0309}$
Fixed multi-hop	Aug	$0.2438_{\pm 0.0384}$	$0.2307_{\pm 0.0376}$	$0.3586_{\pm 0.0513}$	$0.1813_{\pm 0.0337}$
ObsGraph	Q	$0.2346_{\pm 0.0421}$	$0.2209_{\pm 0.0378}$	$0.3567_{\pm 0.0542}$	$0.1658_{\pm 0.0296}$
ObsGraph	Aug	$0.2412_{\pm 0.0369}$	$0.2241_{\pm 0.0353}$	$0.3518_{\pm 0.0487}$	$0.1759_{\pm 0.0324}$
DyNICL	Q	$0.2293_{\pm 0.0468}$	$0.2087_{\pm 0.0394}$	$0.3448_{\pm 0.0571}$	$0.1562_{\pm 0.0308}$
DyNICL	Aug	$0.2387_{\pm 0.0362}$	$0.2176_{\pm 0.0348}$	$0.3481_{\pm 0.0476}$	$0.1728_{\pm 0.0329}$
NetEst	–	$0.3124_{\pm 0.0607}$	$0.3541_{\pm 0.0513}$	$0.5528_{\pm 0.0764}$	$0.2876_{\pm 0.0482}$
TNet	–	$0.3006_{\pm 0.0549}$	$0.3427_{\pm 0.0448}$	$0.5354_{\pm 0.0693}$	$0.2789_{\pm 0.0416}$
Oracle graph	Q	$0.2108_{\pm 0.0337}$	$0.1816_{\pm 0.0293}$	$0.3017_{\pm 0.0438}$	$0.1329_{\pm 0.0226}$
Oracle graph	Aug	$0.2163_{\pm 0.0314}$	$0.1852_{\pm 0.0281}$	$0.2964_{\pm 0.0402}$	$0.1407_{\pm 0.0251}$

Table 5: Effect estimation with 4 neighbors and lag 3.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.3682_{\pm 0.0684}$	$0.4357_{\pm 0.0773}$	$0.5876_{\pm 0.0891}$	$0.3419_{\pm 0.0587}$
Fixed one-hop	Aug	$0.3347_{\pm 0.0716}$	$0.3861_{\pm 0.0908}$	$0.5318_{\pm 0.0946}$	$0.3174_{\pm 0.0712}$
Fixed multi-hop	Q	$0.3071_{\pm 0.0578}$	$0.2916_{\pm 0.0519}$	$0.4213_{\pm 0.0617}$	$0.2384_{\pm 0.0428}$
Fixed multi-hop	Aug	$0.3319_{\pm 0.0623}$	$0.3648_{\pm 0.0967}$	$0.4827_{\pm 0.0904}$	$0.3056_{\pm 0.0659}$
ObsGraph	Q	$0.2918_{\pm 0.0526}$	$0.2683_{\pm 0.0471}$	$0.3894_{\pm 0.0538}$	$0.2171_{\pm 0.0382}$
ObsGraph	Aug	$0.3146_{\pm 0.0557}$	$0.3317_{\pm 0.0836}$	$0.4439_{\pm 0.0761}$	$0.2783_{\pm 0.0574}$
DyNICL	Q	$0.2764_{\pm 0.0549}$	$0.2471_{\pm 0.0458}$	$0.3652_{\pm 0.0483}$	$0.1986_{\pm 0.0371}$
DyNICL	Aug	$0.3048_{\pm 0.0581}$	$0.3162_{\pm 0.0894}$	$0.4226_{\pm 0.0798}$	$0.2627_{\pm 0.0596}$
NetEst	–	$0.3468_{\pm 0.0724}$	$0.4075_{\pm 0.0689}$	$0.6652_{\pm 0.1127}$	$0.3164_{\pm 0.0546}$
TNet	–	$0.3347_{\pm 0.0666}$	$0.4108_{\pm 0.0617}$	$0.6469_{\pm 0.1191}$	$0.3052_{\pm 0.0491}$
Oracle graph	Q	$0.2437_{\pm 0.0406}$	$0.2039_{\pm 0.0334}$	$0.3138_{\pm 0.0397}$	$0.1574_{\pm 0.0268}$
Oracle graph	Aug	$0.2591_{\pm 0.0423}$	$0.2416_{\pm 0.0527}$	$0.3382_{\pm 0.0509}$	$0.1918_{\pm 0.0386}$

Table 6: Effect estimation with 4 neighbors and lag 5.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.5178_{\pm 0.0936}$	$0.6641_{\pm 0.1168}$	$0.8573_{\pm 0.1319}$	$0.5492_{\pm 0.0874}$
Fixed one-hop	Aug	$0.4926_{\pm 0.1197}$	$0.7584_{\pm 0.2876}$	$0.9671_{\pm 0.3028}$	$0.6835_{\pm 0.2061}$
Fixed multi-hop	Q	$0.4117_{\pm 0.0741}$	$0.3538_{\pm 0.0627}$	$0.4639_{\pm 0.0706}$	$0.2847_{\pm 0.0493}$
Fixed multi-hop	Aug	$0.4631_{\pm 0.1048}$	$0.6729_{\pm 0.2537}$	$0.7896_{\pm 0.2641}$	$0.5548_{\pm 0.1736}$
ObsGraph	Q	$0.3942_{\pm 0.0683}$	$0.3176_{\pm 0.0539}$	$0.4187_{\pm 0.0574}$	$0.2531_{\pm 0.0416}$
ObsGraph	Aug	$0.4376_{\pm 0.0917}$	$0.5938_{\pm 0.2114}$	$0.7013_{\pm 0.2186}$	$0.4829_{\pm 0.1428}$
DyNICL	Q	$0.3786_{\pm 0.0709}$	$0.2864_{\pm 0.0498}$	$0.3821_{\pm 0.0487}$	$0.2273_{\pm 0.0359}$
DyNICL	Aug	$0.4219_{\pm 0.0973}$	$0.5527_{\pm 0.2349}$	$0.6498_{\pm 0.2317}$	$0.4486_{\pm 0.1513}$
NetEst	–	$0.4853_{\pm 0.0841}$	$0.4586_{\pm 0.0775}$	$0.6931_{\pm 0.0893}$	$0.3138_{\pm 0.0601}$
TNet	–	$0.4711_{\pm 0.0795}$	$0.4429_{\pm 0.0806}$	$0.6796_{\pm 0.0806}$	$0.3092_{\pm 0.0553}$
Oracle graph	Q	$0.3194_{\pm 0.0496}$	$0.2287_{\pm 0.0361}$	$0.3165_{\pm 0.0408}$	$0.1762_{\pm 0.0274}$
Oracle graph	Aug	$0.3482_{\pm 0.0647}$	$0.3819_{\pm 0.1218}$	$0.4576_{\pm 0.1369}$	$0.3027_{\pm 0.0875}$

Table 7: Effect estimation with 8 neighbors and lag 1.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.3516_{\pm 0.0627}$	$0.4269_{\pm 0.0741}$	$0.6493_{\pm 0.0968}$	$0.3297_{\pm 0.0539}$
Fixed one-hop	Aug	$0.3268_{\pm 0.0513}$	$0.3654_{\pm 0.0547}$	$0.5718_{\pm 0.0682}$	$0.2961_{\pm 0.0456}$
Fixed multi-hop	Q	$0.3187_{\pm 0.0564}$	$0.3628_{\pm 0.0618}$	$0.5624_{\pm 0.0813}$	$0.2806_{\pm 0.0472}$
Fixed multi-hop	Aug	$0.3132_{\pm 0.0461}$	$0.3337_{\pm 0.0439}$	$0.5261_{\pm 0.0584}$	$0.2849_{\pm 0.0413}$
ObsGraph	Q	$0.3094_{\pm 0.0548}$	$0.3416_{\pm 0.0572}$	$0.5378_{\pm 0.0746}$	$0.2631_{\pm 0.0427}$
ObsGraph	Aug	$0.3057_{\pm 0.0426}$	$0.3168_{\pm 0.0381}$	$0.5019_{\pm 0.0517}$	$0.2704_{\pm 0.0378}$
DyNICL	Q	$0.2841_{\pm 0.0607}$	$0.3027_{\pm 0.0593}$	$0.4863_{\pm 0.0864}$	$0.2308_{\pm 0.0441}$
DyNICL	Aug	$0.2986_{\pm 0.0419}$	$0.3079_{\pm 0.0346}$	$0.4747_{\pm 0.0493}$	$0.2536_{\pm 0.0362}$
NetEst	–	$0.3952_{\pm 0.0816}$	$0.4417_{\pm 0.0732}$	$0.6145_{\pm 0.0926}$	$0.3078_{\pm 0.0618}$
TNet	–	$0.3817_{\pm 0.0759}$	$0.4284_{\pm 0.0684}$	$0.5973_{\pm 0.0851}$	$0.3091_{\pm 0.0575}$
Oracle graph	Q	$0.2449_{\pm 0.0378}$	$0.2356_{\pm 0.0384}$	$0.3842_{\pm 0.0529}$	$0.1737_{\pm 0.0286}$
Oracle graph	Aug	$0.2523_{\pm 0.0336}$	$0.2251_{\pm 0.0297}$	$0.3618_{\pm 0.0421}$	$0.1849_{\pm 0.0273}$

Table 8: Effect estimation with 8 neighbors and lag 3.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.4817_{\pm 0.0839}$	$0.6728_{\pm 0.1094}$	$0.8996_{\pm 0.1297}$	$0.5481_{\pm 0.0816}$
Fixed one-hop	Aug	$0.4273_{\pm 0.0958}$	$0.5971_{\pm 0.1573}$	$0.7984_{\pm 0.1649}$	$0.5038_{\pm 0.1142}$
Fixed multi-hop	Q	$0.3642_{\pm 0.0687}$	$0.4493_{\pm 0.0761}$	$0.6217_{\pm 0.0858}$	$0.3694_{\pm 0.0597}$
Fixed multi-hop	Aug	$0.3918_{\pm 0.0834}$	$0.5376_{\pm 0.1398}$	$0.7029_{\pm 0.1326}$	$0.4613_{\pm 0.1028}$
ObsGraph	Q	$0.3386_{\pm 0.0612}$	$0.4037_{\pm 0.0679}$	$0.5591_{\pm 0.0716}$	$0.3258_{\pm 0.0493}$
ObsGraph	Aug	$0.3624_{\pm 0.0727}$	$0.4819_{\pm 0.1174}$	$0.6216_{\pm 0.1073}$	$0.4072_{\pm 0.0869}$
DyNICL	Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
DyNICL	Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
NetEst	–	$0.4706_{\pm 0.0873}$	$0.4905_{\pm 0.0794}$	$0.6638_{\pm 0.1081}$	$0.3267_{\pm 0.0645}$
TNet	–	$0.4733_{\pm 0.0815}$	$0.4743_{\pm 0.0721}$	$0.6463_{\pm 0.1048}$	$0.3146_{\pm 0.0579}$
Oracle graph	Q	$0.2638_{\pm 0.0439}$	$0.2714_{\pm 0.0426}$	$0.3917_{\pm 0.0473}$	$0.2056_{\pm 0.0328}$
Oracle graph	Aug	$0.2816_{\pm 0.0478}$	$0.3109_{\pm 0.0631}$	$0.4235_{\pm 0.0596}$	$0.2468_{\pm 0.0457}$

Table 9: Effect estimation with 8 neighbors and lag 5.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.6584_{\pm 0.1176}$	$0.9631_{\pm 0.1628}$	$1.2746_{\pm 0.1943}$	$0.8269_{\pm 0.1321}$
Fixed one-hop	Aug	$0.6127_{\pm 0.1568}$	$1.1364_{\pm 0.4279}$	$1.5273_{\pm 0.6014}$	$1.0468_{\pm 0.3972}$
Fixed multi-hop	Q	$0.4936_{\pm 0.0847}$	$0.5582_{\pm 0.0903}$	$0.7419_{\pm 0.1096}$	$0.4617_{\pm 0.0748}$
Fixed multi-hop	Aug	$0.5528_{\pm 0.1196}$	$0.9467_{\pm 0.3618}$	$1.2734_{\pm 0.5189}$	$0.8816_{\pm 0.3467}$
ObsGraph	Q	$0.4561_{\pm 0.0729}$	$0.4827_{\pm 0.0738}$	$0.6328_{\pm 0.0861}$	$0.3924_{\pm 0.0586}$
ObsGraph	Aug	$0.5093_{\pm 0.1027}$	$0.8326_{\pm 0.3094}$	$1.1087_{\pm 0.4273}$	$0.7439_{\pm 0.2821}$
DyNICL	Q	$0.4218_{\pm 0.0736}$	$0.4063_{\pm 0.0647}$	$0.5194_{\pm 0.0859}$	$0.3291_{\pm 0.0574}$
DyNICL	Aug	$0.4786_{\pm 0.0984}$	$0.7619_{\pm 0.3376}$	$0.9982_{\pm 0.4938}$	$0.6927_{\pm 0.3659}$
NetEst	–	$0.5491_{\pm 0.0982}$	$0.6083_{\pm 0.0907}$	$0.7754_{\pm 0.1035}$	$0.3982_{\pm 0.0674}$
TNet	–	$0.5326_{\pm 0.0930}$	$0.5926_{\pm 0.0842}$	$0.7530_{\pm 0.0968}$	$0.3843_{\pm 0.0611}$
Oracle graph	Q	$0.3473_{\pm 0.0518}$	$0.3026_{\pm 0.0451}$	$0.4038_{\pm 0.0557}$	$0.2349_{\pm 0.0346}$
Oracle graph	Aug	$0.3819_{\pm 0.0673}$	$0.4798_{\pm 0.1546}$	$0.6217_{\pm 0.2064}$	$0.4028_{\pm 0.1297}$

Table 10: Effect estimation with 12 neighbors and lag 1.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.4518_{\pm 0.0746}$	$0.7827_{\pm 0.0983}$	$1.1794_{\pm 0.1278}$	$0.6183_{\pm 0.0731}$
Fixed one-hop	Aug	$0.4296_{\pm 0.0672}$	$0.5713_{\pm 0.0689}$	$0.8167_{\pm 0.0794}$	$0.5238_{\pm 0.0596}$
Fixed multi-hop	Q	$0.4117_{\pm 0.0618}$	$0.6894_{\pm 0.0796}$	$1.0673_{\pm 0.1037}$	$0.5421_{\pm 0.0593}$
Fixed multi-hop	Aug	$0.4142_{\pm 0.0609}$	$0.5086_{\pm 0.0537}$	$0.7389_{\pm 0.0638}$	$0.4847_{\pm 0.0512}$
ObsGraph	Q	$0.3973_{\pm 0.0574}$	$0.6328_{\pm 0.0659}$	$0.9871_{\pm 0.0846}$	$0.4976_{\pm 0.0478}$
ObsGraph	Aug	$0.4038_{\pm 0.0546}$	$0.4527_{\pm 0.0421}$	$0.6734_{\pm 0.0509}$	$0.4472_{\pm 0.0426}$
DyNICL	Q	$0.3486_{\pm 0.0481}$	$0.5297_{\pm 0.0473}$	$0.8762_{\pm 0.0668}$	$0.4209_{\pm 0.0287}$
DyNICL	Aug	$0.3891_{\pm 0.0597}$	$0.3874_{\pm 0.0358}$	$0.5896_{\pm 0.0427}$	$0.4328_{\pm 0.0493}$
NetEst	–	$0.4208_{\pm 0.0631}$	$0.6845_{\pm 0.0708}$	$1.0947_{\pm 0.1153}$	$0.5853_{\pm 0.0689}$
TNet	–	$0.4072_{\pm 0.0557}$	$0.6671_{\pm 0.0634}$	$1.0710_{\pm 0.1092}$	$0.5689_{\pm 0.0625}$
Oracle graph	Q	$0.2814_{\pm 0.0356}$	$0.3618_{\pm 0.0374}$	$0.5927_{\pm 0.0491}$	$0.2783_{\pm 0.0254}$
Oracle graph	Aug	$0.2976_{\pm 0.0338}$	$0.2849_{\pm 0.0276}$	$0.4381_{\pm 0.0359}$	$0.2667_{\pm 0.0268}$

Table 11: Effect estimation with 12 neighbors and lag 3.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.5829_{\pm 0.0987}$	$1.0483_{\pm 0.1436}$	$1.3817_{\pm 0.1694}$	$0.8416_{\pm 0.1128}$
Fixed one-hop	Aug	$0.5418_{\pm 0.1349}$	$1.0127_{\pm 0.3516}$	$1.2794_{\pm 0.3068}$	$0.8273_{\pm 0.2361}$
Fixed multi-hop	Q	$0.4396_{\pm 0.0738}$	$0.7961_{\pm 0.0927}$	$1.1038_{\pm 0.1124}$	$0.6087_{\pm 0.0726}$
Fixed multi-hop	Aug	$0.4973_{\pm 0.1237}$	$0.9318_{\pm 0.3124}$	$1.1486_{\pm 0.2579}$	$0.7562_{\pm 0.2047}$
ObsGraph	Q	$0.3978_{\pm 0.0617}$	$0.7149_{\pm 0.0663}$	$0.9786_{\pm 0.0798}$	$0.5372_{\pm 0.0489}$
ObsGraph	Aug	$0.4481_{\pm 0.1026}$	$0.8216_{\pm 0.2617}$	$0.9873_{\pm 0.1938}$	$0.6614_{\pm 0.1632}$
DyNICL	Q	$0.3371_{\pm 0.0569}$	$0.5864_{\pm 0.0478}$	$0.8389_{\pm 0.0372}$	$0.4236_{\pm 0.0349}$
DyNICL	Aug	$0.4197_{\pm 0.1194}$	$0.7638_{\pm 0.3371}$	$0.9116_{\pm 0.2183}$	$0.6049_{\pm 0.1867}$
NetEst	–	$0.4816_{\pm 0.0749}$	$0.7208_{\pm 0.0646}$	$1.0976_{\pm 0.0962}$	$0.6124_{\pm 0.0728}$
TNet	–	$0.4651_{\pm 0.0687}$	$0.7033_{\pm 0.0571}$	$1.1003_{\pm 0.0886}$	$0.5965_{\pm 0.0667}$
Oracle graph	Q	$0.2789_{\pm 0.0386}$	$0.3987_{\pm 0.0381}$	$0.5714_{\pm 0.0437}$	$0.2846_{\pm 0.0283}$
Oracle graph	Aug	$0.3124_{\pm 0.0517}$	$0.4621_{\pm 0.0874}$	$0.6048_{\pm 0.0786}$	$0.3527_{\pm 0.0619}$

Table 12: Effect estimation with 12 neighbors and lag 5.

Structure	Scheme	IDE	IPE	ITE	INT
Fixed one-hop	Q	$0.7613_{\pm 0.1328}$	$1.3379_{\pm 0.1946}$	$1.7462_{\pm 0.2387}$	$1.1028_{\pm 0.1563}$
Fixed one-hop	Aug	$0.7496_{\pm 0.2217}$	$1.6893_{\pm 0.6574}$	$2.2186_{\pm 0.8093}$	$1.4537_{\pm 0.5218}$
Fixed multi-hop	Q	$0.5497_{\pm 0.0896}$	$0.9428_{\pm 0.1214}$	$1.2783_{\pm 0.1469}$	$0.7391_{\pm 0.0982}$
Fixed multi-hop	Aug	$0.6774_{\pm 0.1873}$	$1.4538_{\pm 0.5536}$	$1.9297_{\pm 0.6841}$	$1.2176_{\pm 0.4189}$
ObsGraph	Q	$0.4893_{\pm 0.0708}$	$0.8261_{\pm 0.0897}$	$1.0974_{\pm 0.1036}$	$0.6238_{\pm 0.0714}$
ObsGraph	Aug	$0.5987_{\pm 0.1524}$	$1.2849_{\pm 0.4683}$	$1.6928_{\pm 0.5467}$	$1.0394_{\pm 0.3238}$
DyNICL	Q	$0.4268_{\pm 0.0589}$	$0.6763_{\pm 0.0718}$	$0.9197_{\pm 0.0924}$	$0.4931_{\pm 0.0762}$
DyNICL	Aug	$0.5639_{\pm 0.1686}$	$1.1736_{\pm 0.5247}$	$1.5284_{\pm 0.5981}$	$0.9478_{\pm 0.3369}$
NetEst	–	$0.6507_{\pm 0.0793}$	$0.8462_{\pm 0.0751}$	$1.1432_{\pm 0.0824}$	$0.6895_{\pm 0.0663}$
TNet	–	$0.6319_{\pm 0.0705}$	$0.8254_{\pm 0.0683}$	$1.1160_{\pm 0.0689}$	$0.6711_{\pm 0.0602}$
Oracle graph	Q	$0.3517_{\pm 0.0453}$	$0.4528_{\pm 0.0486}$	$0.6289_{\pm 0.0617}$	$0.3296_{\pm 0.0378}$
Oracle graph	Aug	$0.4138_{\pm 0.0819}$	$0.7164_{\pm 0.2187}$	$0.9463_{\pm 0.2874}$	$0.5841_{\pm 0.1646}$

E.2 Module ablations

Table 13: Module ablations with 4 neighbors and lag 1.

Variant	Scheme	IDE	IPE	ITE	INT
DyNICL	Q	$0.2293_{\pm 0.0468}$	$0.2087_{\pm 0.0394}$	$0.3448_{\pm 0.0571}$	$0.1562_{\pm 0.0308}$
DyNICL	Aug	$0.2387_{\pm 0.0362}$	$0.2176_{\pm 0.0348}$	$0.3481_{\pm 0.0476}$	$0.1728_{\pm 0.0329}$
w/o peer pathway	Q	$0.2426_{\pm 0.0427}$	$0.3694_{\pm 0.0618}$	$0.5063_{\pm 0.0739}$	$0.2847_{\pm 0.0472}$
w/o peer pathway	Aug	$0.2479_{\pm 0.0401}$	$0.3168_{\pm 0.0583}$	$0.4537_{\pm 0.0684}$	$0.2581_{\pm 0.0456}$
w/o temporal attention	Q	$0.2328_{\pm 0.0451}$	$0.2119_{\pm 0.0386}$	$0.3496_{\pm 0.0558}$	$0.1594_{\pm 0.0312}$
w/o temporal attention	Aug	$0.2416_{\pm 0.0368}$	$0.2213_{\pm 0.0357}$	$0.3527_{\pm 0.0489}$	$0.1771_{\pm 0.0334}$
w/o peer attention	Q	$0.2347_{\pm 0.0429}$	$0.2246_{\pm 0.0407}$	$0.3618_{\pm 0.0584}$	$0.1693_{\pm 0.0317}$
w/o peer attention	Aug	$0.2421_{\pm 0.0356}$	$0.2298_{\pm 0.0363}$	$0.3612_{\pm 0.0491}$	$0.1846_{\pm 0.0348}$
w/o token mixer	Q	$0.2271_{\pm 0.0416}$	$0.2138_{\pm 0.0381}$	$0.3462_{\pm 0.0527}$	$0.1587_{\pm 0.0294}$
w/o token mixer	Aug	$0.2369_{\pm 0.0347}$	$0.2231_{\pm 0.0349}$	$0.3506_{\pm 0.0458}$	$0.1743_{\pm 0.0316}$

Table 14: Module ablations with 4 neighbors and lag 5.

Variant	Scheme	IDE	IPE	ITE	INT
DyNICL	Q	$0.3786_{\pm 0.0709}$	$0.2864_{\pm 0.0498}$	$0.3821_{\pm 0.0487}$	$0.2273_{\pm 0.0359}$
DyNICL	Aug	$0.4219_{\pm 0.0973}$	$0.5527_{\pm 0.2349}$	$0.6498_{\pm 0.2317}$	$0.4486_{\pm 0.1513}$
w/o peer pathway	Q	$0.4097_{\pm 0.0736}$	$0.5783_{\pm 0.0907}$	$0.7164_{\pm 0.1028}$	$0.4918_{\pm 0.0724}$
w/o peer pathway	Aug	$0.4516_{\pm 0.1094}$	$0.7138_{\pm 0.2786}$	$0.8917_{\pm 0.3092}$	$0.6329_{\pm 0.2037}$
w/o temporal attention	Q	$0.4672_{\pm 0.0813}$	$0.3218_{\pm 0.0537}$	$0.5369_{\pm 0.0734}$	$0.3386_{\pm 0.0501}$
w/o temporal attention	Aug	$0.4873_{\pm 0.1126}$	$0.5894_{\pm 0.2461}$	$0.7716_{\pm 0.2684}$	$0.5437_{\pm 0.1748}$
w/o peer attention	Q	$0.3859_{\pm 0.0687}$	$0.3087_{\pm 0.0516}$	$0.4073_{\pm 0.0528}$	$0.2461_{\pm 0.0374}$
w/o peer attention	Aug	$0.4268_{\pm 0.0936}$	$0.5781_{\pm 0.2297}$	$0.6794_{\pm 0.2381}$	$0.4713_{\pm 0.1559}$
w/o token mixer	Q	$0.3984_{\pm 0.0721}$	$0.3029_{\pm 0.0504}$	$0.4196_{\pm 0.0563}$	$0.2537_{\pm 0.0388}$
w/o token mixer	Aug	$0.4387_{\pm 0.0982}$	$0.5716_{\pm 0.2418}$	$0.6913_{\pm 0.2506}$	$0.4798_{\pm 0.1594}$

Table 15: Module ablations with 8 neighbors and lag 3.

Variant	Scheme	IDE	IPE	ITE	INT
DyNICL	Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
DyNICL	Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
w/o peer pathway	Q	$0.3387_{\pm 0.0718}$	$0.7296_{\pm 0.1184}$	$0.9172_{\pm 0.1396}$	$0.6148_{\pm 0.0917}$
w/o peer pathway	Aug	$0.3574_{\pm 0.0849}$	$0.6341_{\pm 0.1728}$	$0.8093_{\pm 0.1847}$	$0.5496_{\pm 0.1321}$
w/o temporal attention	Q	$0.3498_{\pm 0.0746}$	$0.3837_{\pm 0.0719}$	$0.5861_{\pm 0.0874}$	$0.3452_{\pm 0.0583}$
w/o temporal attention	Aug	$0.3762_{\pm 0.0827}$	$0.4659_{\pm 0.1384}$	$0.6428_{\pm 0.1286}$	$0.4241_{\pm 0.1047}$
w/o peer attention	Q	$0.3196_{\pm 0.0658}$	$0.4068_{\pm 0.0731}$	$0.5574_{\pm 0.0819}$	$0.3291_{\pm 0.0546}$
w/o peer attention	Aug	$0.3518_{\pm 0.0753}$	$0.4763_{\pm 0.1287}$	$0.6119_{\pm 0.1174}$	$0.4097_{\pm 0.0958}$
w/o token mixer	Q	$0.3147_{\pm 0.0672}$	$0.3709_{\pm 0.0684}$	$0.5296_{\pm 0.0763}$	$0.3078_{\pm 0.0517}$
w/o token mixer	Aug	$0.3476_{\pm 0.0741}$	$0.4492_{\pm 0.1263}$	$0.5867_{\pm 0.1138}$	$0.3884_{\pm 0.0936}$

Table 16: Module ablations with 12 neighbors and lag 1.

Variant	Scheme	IDE	IPE	ITE	INT
DyNICL	Q	$0.3486_{\pm 0.0481}$	$0.5297_{\pm 0.0473}$	$0.8762_{\pm 0.0668}$	$0.4209_{\pm 0.0287}$
DyNICL	Aug	$0.3891_{\pm 0.0597}$	$0.3874_{\pm 0.0358}$	$0.5896_{\pm 0.0427}$	$0.4328_{\pm 0.0493}$
w/o peer pathway	Q	$0.4027_{\pm 0.0576}$	$0.9843_{\pm 0.1079}$	$1.3976_{\pm 0.1384}$	$0.8271_{\pm 0.0768}$
w/o peer pathway	Aug	$0.4316_{\pm 0.0693}$	$0.6839_{\pm 0.0861}$	$0.9748_{\pm 0.1067}$	$0.7083_{\pm 0.0834}$
w/o temporal attention	Q	$0.3519_{\pm 0.0467}$	$0.5358_{\pm 0.0489}$	$0.8847_{\pm 0.0641}$	$0.4273_{\pm 0.0296}$
w/o temporal attention	Aug	$0.3918_{\pm 0.0582}$	$0.3916_{\pm 0.0367}$	$0.5952_{\pm 0.0439}$	$0.4381_{\pm 0.0487}$
w/o peer attention	Q	$0.3774_{\pm 0.0459}$	$0.6816_{\pm 0.0634}$	$1.0638_{\pm 0.0817}$	$0.5487_{\pm 0.0412}$
w/o peer attention	Aug	$0.4107_{\pm 0.0568}$	$0.4913_{\pm 0.0449}$	$0.7318_{\pm 0.0586}$	$0.5394_{\pm 0.0531}$
w/o token mixer	Q	$0.3571_{\pm 0.0476}$	$0.5489_{\pm 0.0481}$	$0.8993_{\pm 0.0662}$	$0.4378_{\pm 0.0301}$
w/o token mixer	Aug	$0.3936_{\pm 0.0574}$	$0.3998_{\pm 0.0372}$	$0.6071_{\pm 0.0453}$	$0.4467_{\pm 0.0498}$

Table 17: Module ablations with 12 neighbors and lag 5.

Variant	Scheme	IDE	IPE	ITE	INT
DyNICL	Q	$0.4268_{\pm 0.0589}$	$0.6763_{\pm 0.0718}$	$0.9197_{\pm 0.0924}$	$0.4931_{\pm 0.0762}$
DyNICL	Aug	$0.5639_{\pm 0.1686}$	$1.1736_{\pm 0.5247}$	$1.5284_{\pm 0.5981}$	$0.9478_{\pm 0.3369}$
w/o peer pathway	Q	$0.5074_{\pm 0.0761}$	$1.3278_{\pm 0.1693}$	$1.7469_{\pm 0.2148}$	$1.0986_{\pm 0.1427}$
w/o peer pathway	Aug	$0.6518_{\pm 0.2047}$	$1.7293_{\pm 0.6814}$	$2.2671_{\pm 0.8476}$	$1.4689_{\pm 0.5143}$
w/o temporal attention	Q	$0.5481_{\pm 0.0817}$	$0.7429_{\pm 0.0836}$	$1.1634_{\pm 0.1268}$	$0.6817_{\pm 0.0914}$
w/o temporal attention	Aug	$0.6562_{\pm 0.1928}$	$1.2617_{\pm 0.5493}$	$1.7938_{\pm 0.6841}$	$1.1526_{\pm 0.3927}$
w/o peer attention	Q	$0.4619_{\pm 0.0607}$	$0.8834_{\pm 0.1028}$	$1.1976_{\pm 0.1391}$	$0.7028_{\pm 0.0863}$
w/o peer attention	Aug	$0.5937_{\pm 0.1619}$	$1.3816_{\pm 0.5638}$	$1.8462_{\pm 0.6739}$	$1.1934_{\pm 0.3748}$
w/o token mixer	Q	$0.4936_{\pm 0.0694}$	$0.7812_{\pm 0.0871}$	$1.1248_{\pm 0.1176}$	$0.6593_{\pm 0.0837}$
w/o token mixer	Aug	$0.6194_{\pm 0.1762}$	$1.2789_{\pm 0.5364}$	$1.7426_{\pm 0.6528}$	$1.1217_{\pm 0.3681}$

E.3 HYPERPARAMETER SENSITIVITY

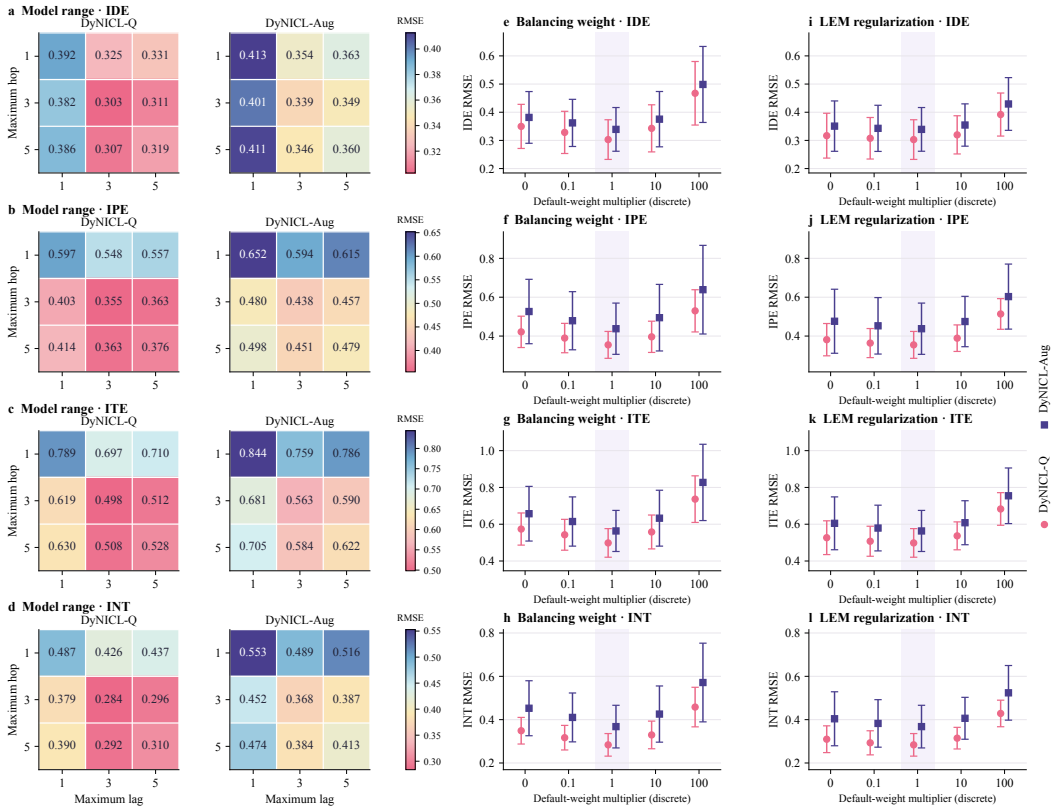


Figure 5: Hyperparameter sensitivity for all four effects on the 8×3 setting. Left: maximum hop and lag, with a shared color scale across Q and Aug within each effect. Middle and right: balancing and LEM regularization weights relative to their defaults. Pale shading marks the lowest observed mean in each weight sweep. The intervention window remains fixed.

Table 18: Model range on 8×3 data: maximum hop / lag, with a fixed intervention window.

Level	Method	IDE	IPE	ITE	INT
1 / 1	Q	$0.3917_{\pm 0.0836}$	$0.5968_{\pm 0.0974}$	$0.7894_{\pm 0.1126}$	$0.4873_{\pm 0.0769}$
1 / 1	Aug	$0.4126_{\pm 0.0987}$	$0.6519_{\pm 0.1723}$	$0.8437_{\pm 0.1638}$	$0.5528_{\pm 0.1361}$
1 / 3	Q	$0.3248_{\pm 0.0741}$	$0.5483_{\pm 0.0896}$	$0.6971_{\pm 0.0984}$	$0.4257_{\pm 0.0678}$
1 / 3	Aug	$0.3537_{\pm 0.0862}$	$0.5936_{\pm 0.1547}$	$0.7586_{\pm 0.1489}$	$0.4891_{\pm 0.1214}$
1 / 5	Q	$0.3306_{\pm 0.0807}$	$0.5569_{\pm 0.0961}$	$0.7098_{\pm 0.1057}$	$0.4374_{\pm 0.0736}$
1 / 5	Aug	$0.3628_{\pm 0.0943}$	$0.6147_{\pm 0.1816}$	$0.7862_{\pm 0.1714}$	$0.5163_{\pm 0.1439}$
3 / 1	Q	$0.3824_{\pm 0.0798}$	$0.4027_{\pm 0.0746}$	$0.6187_{\pm 0.0873}$	$0.3789_{\pm 0.0631}$
3 / 1	Aug	$0.4013_{\pm 0.0916}$	$0.4798_{\pm 0.1427}$	$0.6814_{\pm 0.1326}$	$0.4516_{\pm 0.1098}$
3 / 3	Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
3 / 3	Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
3 / 5	Q	$0.3107_{\pm 0.0759}$	$0.3631_{\pm 0.0728}$	$0.5119_{\pm 0.0836}$	$0.2956_{\pm 0.0567}$
3 / 5	Aug	$0.3486_{\pm 0.0847}$	$0.4572_{\pm 0.1483}$	$0.5897_{\pm 0.1268}$	$0.3874_{\pm 0.1116}$
5 / 1	Q	$0.3861_{\pm 0.0854}$	$0.4138_{\pm 0.0809}$	$0.6296_{\pm 0.0951}$	$0.3897_{\pm 0.0682}$
5 / 1	Aug	$0.4108_{\pm 0.1023}$	$0.4976_{\pm 0.1658}$	$0.7048_{\pm 0.1493}$	$0.4741_{\pm 0.1273}$
5 / 3	Q	$0.3074_{\pm 0.0738}$	$0.3627_{\pm 0.0771}$	$0.5076_{\pm 0.0849}$	$0.2918_{\pm 0.0586}$
5 / 3	Aug	$0.3462_{\pm 0.0859}$	$0.4513_{\pm 0.1517}$	$0.5841_{\pm 0.1297}$	$0.3836_{\pm 0.1132}$
5 / 5	Q	$0.3186_{\pm 0.0821}$	$0.3764_{\pm 0.0853}$	$0.5284_{\pm 0.0937}$	$0.3097_{\pm 0.0648}$
5 / 5	Aug	$0.3598_{\pm 0.0976}$	$0.4791_{\pm 0.1794}$	$0.6218_{\pm 0.1532}$	$0.4128_{\pm 0.1357}$

Table 19: Balancing weight on the 8×3 setting. Levels are multipliers of the default coefficient.

Level	Method	IDE	IPE	ITE	INT
0×	Q	$0.3497_{\pm 0.0783}$	$0.4216_{\pm 0.0807}$	$0.5738_{\pm 0.0874}$	$0.3489_{\pm 0.0614}$
0×	Aug	$0.3816_{\pm 0.0918}$	$0.5263_{\pm 0.1657}$	$0.6571_{\pm 0.1486}$	$0.4527_{\pm 0.1269}$
0.1×	Q	$0.3284_{\pm 0.0749}$	$0.3897_{\pm 0.0756}$	$0.5426_{\pm 0.0839}$	$0.3168_{\pm 0.0567}$
0.1×	Aug	$0.3621_{\pm 0.0837}$	$0.4789_{\pm 0.1496}$	$0.6147_{\pm 0.1338}$	$0.4103_{\pm 0.1128}$
1×	Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
1×	Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
10×	Q	$0.3428_{\pm 0.0836}$	$0.3961_{\pm 0.0804}$	$0.5579_{\pm 0.0921}$	$0.3295_{\pm 0.0638}$
10×	Aug	$0.3754_{\pm 0.0982}$	$0.4946_{\pm 0.1713}$	$0.6328_{\pm 0.1519}$	$0.4258_{\pm 0.1297}$
100×	Q	$0.4673_{\pm 0.1128}$	$0.5298_{\pm 0.1086}$	$0.7364_{\pm 0.1267}$	$0.4581_{\pm 0.0913}$
100×	Aug	$0.4986_{\pm 0.1349}$	$0.6387_{\pm 0.2281}$	$0.8273_{\pm 0.2074}$	$0.5714_{\pm 0.1816}$

Table 20: LEM regularization on the 8×3 setting. Levels are multipliers of the default coefficient.

Level	Method	IDE	IPE	ITE	INT
0×	Q	$0.3168_{\pm 0.0796}$	$0.3814_{\pm 0.0831}$	$0.5267_{\pm 0.0918}$	$0.3097_{\pm 0.0618}$
0×	Aug	$0.3507_{\pm 0.0894}$	$0.4762_{\pm 0.1648}$	$0.6049_{\pm 0.1437}$	$0.4041_{\pm 0.1246}$
0.1×	Q	$0.3076_{\pm 0.0737}$	$0.3639_{\pm 0.0748}$	$0.5074_{\pm 0.0816}$	$0.2931_{\pm 0.0556}$
0.1×	Aug	$0.3428_{\pm 0.0819}$	$0.4527_{\pm 0.1453}$	$0.5792_{\pm 0.1243}$	$0.3826_{\pm 0.1097}$
1×	Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
1×	Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
10×	Q	$0.3197_{\pm 0.0678}$	$0.3891_{\pm 0.0683}$	$0.5368_{\pm 0.0759}$	$0.3146_{\pm 0.0499}$
10×	Aug	$0.3546_{\pm 0.0751}$	$0.4748_{\pm 0.1296}$	$0.6081_{\pm 0.1196}$	$0.4063_{\pm 0.0967}$
100×	Q	$0.3918_{\pm 0.0764}$	$0.5137_{\pm 0.0789}$	$0.6829_{\pm 0.0887}$	$0.4286_{\pm 0.0613}$
100×	Aug	$0.4293_{\pm 0.0936}$	$0.6031_{\pm 0.1674}$	$0.7546_{\pm 0.1512}$	$0.5238_{\pm 0.1261}$

E.4 Graph error and omitted information

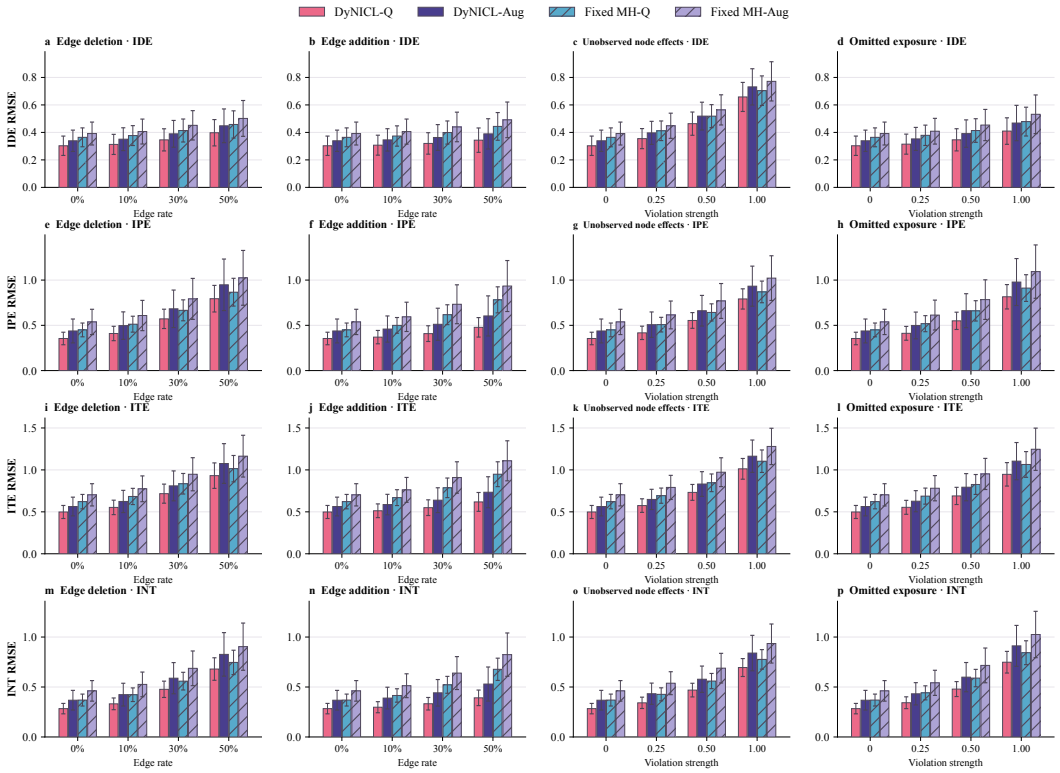


Figure 6: Robustness for all four effects on the 8×3 setting. Columns vary edge deletion, edge addition, unobserved node effects, and omitted exposure; rows show IDE, IPE, ITE, and INT. Hatched bars denote fixed multi-hop exposure. Graph perturbations preserve the data and causal targets, whereas each node-effect and omitted-exposure level uses its corresponding counterfactual truth.

Table 21: Edge deletion on the 8×3 setting. Levels use the original edge count; truth remains fixed.

Level	Method	IDE	IPE	ITE	INT
0%	DyNICL-Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
0%	DyNICL-Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
0%	Fixed MH-Q	$0.3642_{\pm 0.0687}$	$0.4493_{\pm 0.0761}$	$0.6217_{\pm 0.0858}$	$0.3694_{\pm 0.0597}$
0%	Fixed MH-Aug	$0.3918_{\pm 0.0834}$	$0.5376_{\pm 0.1398}$	$0.7029_{\pm 0.1326}$	$0.4613_{\pm 0.1028}$
10%	DyNICL-Q	$0.3127_{\pm 0.0731}$	$0.4098_{\pm 0.0794}$	$0.5529_{\pm 0.0861}$	$0.3316_{\pm 0.0587}$
10%	DyNICL-Aug	$0.3508_{\pm 0.0829}$	$0.4971_{\pm 0.1516}$	$0.6238_{\pm 0.1324}$	$0.4237_{\pm 0.1138}$
10%	Fixed MH-Q	$0.3769_{\pm 0.0724}$	$0.5127_{\pm 0.0879}$	$0.6834_{\pm 0.0967}$	$0.4228_{\pm 0.0681}$
10%	Fixed MH-Aug	$0.4063_{\pm 0.0907}$	$0.6084_{\pm 0.1673}$	$0.7746_{\pm 0.1538}$	$0.5261_{\pm 0.1234}$
30%	DyNICL-Q	$0.3458_{\pm 0.0806}$	$0.5714_{\pm 0.1068}$	$0.7168_{\pm 0.1137}$	$0.4769_{\pm 0.0817}$
30%	DyNICL-Aug	$0.3897_{\pm 0.0974}$	$0.6826_{\pm 0.2071}$	$0.8097_{\pm 0.1783}$	$0.5883_{\pm 0.1559}$
30%	Fixed MH-Q	$0.4138_{\pm 0.0836}$	$0.6659_{\pm 0.1157}$	$0.8349_{\pm 0.1226}$	$0.5587_{\pm 0.0884}$
30%	Fixed MH-Aug	$0.4516_{\pm 0.1068}$	$0.7921_{\pm 0.2264}$	$0.9481_{\pm 0.1974}$	$0.6869_{\pm 0.1736}$
50%	DyNICL-Q	$0.3976_{\pm 0.0958}$	$0.7943_{\pm 0.1469}$	$0.9316_{\pm 0.1519}$	$0.6798_{\pm 0.1126}$
50%	DyNICL-Aug	$0.4481_{\pm 0.1227}$	$0.9478_{\pm 0.2846}$	$1.0762_{\pm 0.2368}$	$0.8264_{\pm 0.2178}$
50%	Fixed MH-Q	$0.4579_{\pm 0.0991}$	$0.8662_{\pm 0.1534}$	$1.0147_{\pm 0.1573}$	$0.7461_{\pm 0.1219}$
50%	Fixed MH-Aug	$0.5018_{\pm 0.1304}$	$1.0247_{\pm 0.3028}$	$1.1639_{\pm 0.2491}$	$0.9037_{\pm 0.2361}$

Table 22: Edge addition on the 8×3 setting. Levels use the original edge count; truth remains fixed.

Level	Method	IDE	IPE	ITE	INT
0%	DyNICL-Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
0%	DyNICL-Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
0%	Fixed MH-Q	$0.3642_{\pm 0.0687}$	$0.4493_{\pm 0.0761}$	$0.6217_{\pm 0.0858}$	$0.3694_{\pm 0.0597}$
0%	Fixed MH-Aug	$0.3918_{\pm 0.0834}$	$0.5376_{\pm 0.1398}$	$0.7029_{\pm 0.1326}$	$0.4613_{\pm 0.1028}$
10%	DyNICL-Q	$0.3071_{\pm 0.0726}$	$0.3697_{\pm 0.0738}$	$0.5128_{\pm 0.0819}$	$0.2978_{\pm 0.0554}$
10%	DyNICL-Aug	$0.3456_{\pm 0.0814}$	$0.4589_{\pm 0.1447}$	$0.5867_{\pm 0.1221}$	$0.3892_{\pm 0.1086}$
10%	Fixed MH-Q	$0.3738_{\pm 0.0739}$	$0.4996_{\pm 0.0867}$	$0.6693_{\pm 0.0946}$	$0.4153_{\pm 0.0671}$
10%	Fixed MH-Aug	$0.4059_{\pm 0.0913}$	$0.5937_{\pm 0.1618}$	$0.7618_{\pm 0.1483}$	$0.5128_{\pm 0.1194}$
30%	DyNICL-Q	$0.3198_{\pm 0.0781}$	$0.4083_{\pm 0.0869}$	$0.5497_{\pm 0.0936}$	$0.3326_{\pm 0.0627}$
30%	DyNICL-Aug	$0.3631_{\pm 0.0938}$	$0.5116_{\pm 0.1774}$	$0.6391_{\pm 0.1479}$	$0.4419_{\pm 0.1328}$
30%	Fixed MH-Q	$0.3991_{\pm 0.0842}$	$0.6174_{\pm 0.1098}$	$0.7864_{\pm 0.1168}$	$0.5218_{\pm 0.0846}$
30%	Fixed MH-Aug	$0.4397_{\pm 0.1076}$	$0.7328_{\pm 0.2149}$	$0.9096_{\pm 0.1872}$	$0.6394_{\pm 0.1647}$
50%	DyNICL-Q	$0.3436_{\pm 0.0887}$	$0.4779_{\pm 0.1076}$	$0.6183_{\pm 0.1124}$	$0.3917_{\pm 0.0778}$
50%	DyNICL-Aug	$0.3894_{\pm 0.1102}$	$0.6038_{\pm 0.2216}$	$0.7338_{\pm 0.1846}$	$0.5297_{\pm 0.1704}$
50%	Fixed MH-Q	$0.4437_{\pm 0.1008}$	$0.7816_{\pm 0.1439}$	$0.9487_{\pm 0.1481}$	$0.6773_{\pm 0.1112}$
50%	Fixed MH-Aug	$0.4913_{\pm 0.1297}$	$0.9342_{\pm 0.2817}$	$1.1084_{\pm 0.2387}$	$0.8239_{\pm 0.2168}$

E.5 Assignment support

Table 23: Unobserved node effects on the 8×3 setting. The common assignment/outcome effect is fixed within each design and withheld from the learner. Each level has its corresponding counterfactual truth.

Level	Method	IDE	IPE	ITE	INT
0	DyNICL-Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
0	DyNICL-Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
0	Fixed MH-Q	$0.3642_{\pm 0.0687}$	$0.4493_{\pm 0.0761}$	$0.6217_{\pm 0.0858}$	$0.3694_{\pm 0.0597}$
0	Fixed MH-Aug	$0.3918_{\pm 0.0834}$	$0.5376_{\pm 0.1398}$	$0.7029_{\pm 0.1326}$	$0.4613_{\pm 0.1028}$
0.25	DyNICL-Q	$0.3548_{\pm 0.0731}$	$0.4169_{\pm 0.0737}$	$0.5746_{\pm 0.0809}$	$0.3416_{\pm 0.0568}$
0.25	DyNICL-Aug	$0.3967_{\pm 0.0839}$	$0.5072_{\pm 0.1406}$	$0.6481_{\pm 0.1206}$	$0.4328_{\pm 0.1063}$
0.25	Fixed MH-Q	$0.4119_{\pm 0.0714}$	$0.5076_{\pm 0.0809}$	$0.6928_{\pm 0.0907}$	$0.4267_{\pm 0.0652}$
0.25	Fixed MH-Aug	$0.4483_{\pm 0.0916}$	$0.6149_{\pm 0.1537}$	$0.7913_{\pm 0.1438}$	$0.5381_{\pm 0.1148}$
0.50	DyNICL-Q	$0.4637_{\pm 0.0846}$	$0.5538_{\pm 0.0869}$	$0.7319_{\pm 0.0947}$	$0.4691_{\pm 0.0683}$
0.50	DyNICL-Aug	$0.5186_{\pm 0.1012}$	$0.6614_{\pm 0.1693}$	$0.8317_{\pm 0.1472}$	$0.5779_{\pm 0.1316}$
0.50	Fixed MH-Q	$0.5184_{\pm 0.0837}$	$0.6427_{\pm 0.0948}$	$0.8461_{\pm 0.1054}$	$0.5596_{\pm 0.0761}$
0.50	Fixed MH-Aug	$0.5639_{\pm 0.1098}$	$0.7693_{\pm 0.1926}$	$0.9726_{\pm 0.1719}$	$0.6887_{\pm 0.1482}$
1.00	DyNICL-Q	$0.6581_{\pm 0.1057}$	$0.7916_{\pm 0.1118}$	$1.0128_{\pm 0.1236}$	$0.6948_{\pm 0.0897}$
1.00	DyNICL-Aug	$0.7314_{\pm 0.1319}$	$0.9327_{\pm 0.2214}$	$1.1634_{\pm 0.1928}$	$0.8396_{\pm 0.1773}$
1.00	Fixed MH-Q	$0.7038_{\pm 0.1074}$	$0.8692_{\pm 0.1196}$	$1.1037_{\pm 0.1341}$	$0.7763_{\pm 0.0978}$
1.00	Fixed MH-Aug	$0.7716_{\pm 0.1427}$	$1.0218_{\pm 0.2469}$	$1.2798_{\pm 0.2167}$	$0.9351_{\pm 0.1954}$

Table 24: Omitted exposure on 8×3 data, with level-specific counterfactual truth.

Level	Method	IDE	IPE	ITE	INT
0	DyNICL-Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
0	DyNICL-Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
0	Fixed MH-Q	$0.3642_{\pm 0.0687}$	$0.4493_{\pm 0.0761}$	$0.6217_{\pm 0.0858}$	$0.3694_{\pm 0.0597}$
0	Fixed MH-Aug	$0.3918_{\pm 0.0834}$	$0.5376_{\pm 0.1398}$	$0.7029_{\pm 0.1326}$	$0.4613_{\pm 0.1028}$
0.25	DyNICL-Q	$0.3146_{\pm 0.0729}$	$0.4118_{\pm 0.0763}$	$0.5538_{\pm 0.0831}$	$0.3427_{\pm 0.0591}$
0.25	DyNICL-Aug	$0.3528_{\pm 0.0837}$	$0.4986_{\pm 0.1478}$	$0.6269_{\pm 0.1257}$	$0.4321_{\pm 0.1114}$
0.25	Fixed MH-Q	$0.3789_{\pm 0.0741}$	$0.5183_{\pm 0.0887}$	$0.6874_{\pm 0.0958}$	$0.4416_{\pm 0.0702}$
0.25	Fixed MH-Aug	$0.4087_{\pm 0.0936}$	$0.6134_{\pm 0.1652}$	$0.7816_{\pm 0.1504}$	$0.5418_{\pm 0.1257}$
0.50	DyNICL-Q	$0.3461_{\pm 0.0807}$	$0.5497_{\pm 0.0954}$	$0.6897_{\pm 0.1026}$	$0.4793_{\pm 0.0748}$
0.50	DyNICL-Aug	$0.3914_{\pm 0.0998}$	$0.6619_{\pm 0.1876}$	$0.7938_{\pm 0.1619}$	$0.5987_{\pm 0.1463}$
0.50	Fixed MH-Q	$0.4137_{\pm 0.0859}$	$0.6598_{\pm 0.1123}$	$0.8263_{\pm 0.1187}$	$0.5894_{\pm 0.0871}$
0.50	Fixed MH-Aug	$0.4536_{\pm 0.1137}$	$0.7841_{\pm 0.2189}$	$0.9517_{\pm 0.1856}$	$0.7162_{\pm 0.1738}$
1.00	DyNICL-Q	$0.4098_{\pm 0.0964}$	$0.8156_{\pm 0.1347}$	$0.9471_{\pm 0.1398}$	$0.7481_{\pm 0.1083}$
1.00	DyNICL-Aug	$0.4683_{\pm 0.1286}$	$0.9784_{\pm 0.2593}$	$1.1046_{\pm 0.2217}$	$0.9117_{\pm 0.2046}$
1.00	Fixed MH-Q	$0.4786_{\pm 0.1048}$	$0.9117_{\pm 0.1476}$	$1.0649_{\pm 0.1532}$	$0.8429_{\pm 0.1197}$
1.00	Fixed MH-Aug	$0.5319_{\pm 0.1407}$	$1.0936_{\pm 0.2948}$	$1.2463_{\pm 0.2518}$	$1.0254_{\pm 0.2319}$

Table 25: Assignment-support sensitivity with 8 neighbors and lag 3. ϵ is the data-generating probability floor; estimated propensities are clipped to $[0.01, 0.99]$ at every level.

ϵ	Scheme	IDE	IPE	ITE	INT
0.2000	Q	$0.2908_{\pm 0.0617}$	$0.3371_{\pm 0.0589}$	$0.4761_{\pm 0.0679}$	$0.2694_{\pm 0.0446}$
0.2000	Aug	$0.2819_{\pm 0.0543}$	$0.3158_{\pm 0.0526}$	$0.4587_{\pm 0.0634}$	$0.2517_{\pm 0.0418}$
0.1000	Q	$0.3029_{\pm 0.0704}$	$0.3546_{\pm 0.0692}$	$0.4983_{\pm 0.0778}$	$0.2837_{\pm 0.0521}$
0.1000	Aug	$0.3391_{\pm 0.0776}$	$0.4378_{\pm 0.1319}$	$0.5634_{\pm 0.1117}$	$0.3679_{\pm 0.0984}$
0.0300	Q	$0.3267_{\pm 0.0813}$	$0.3978_{\pm 0.0837}$	$0.5498_{\pm 0.0916}$	$0.3216_{\pm 0.0659}$
0.0300	Aug	$0.4168_{\pm 0.1264}$	$0.6481_{\pm 0.2783}$	$0.7829_{\pm 0.2473}$	$0.5529_{\pm 0.2076}$
0.0050	Q	$0.3714_{\pm 0.1058}$	$0.4769_{\pm 0.1156}$	$0.6387_{\pm 0.1248}$	$0.3947_{\pm 0.0891}$
0.0050	Aug	$0.5873_{\pm 0.2449}$	$1.0936_{\pm 0.6487}$	$1.2641_{\pm 0.5869}$	$0.9384_{\pm 0.4918}$

Table 26: Assignment-support sensitivity with 12 neighbors and lag 5. ϵ is the data-generating probability floor; estimated propensities are clipped to $[0.01, 0.99]$ at every level.

ϵ	Scheme	IDE	IPE	ITE	INT
0.2000	Q	$0.4019_{\pm 0.0526}$	$0.6238_{\pm 0.0621}$	$0.8536_{\pm 0.0817}$	$0.4487_{\pm 0.0619}$
0.2000	Aug	$0.3926_{\pm 0.0498}$	$0.5817_{\pm 0.0594}$	$0.8194_{\pm 0.0786}$	$0.4213_{\pm 0.0576}$
0.1000	Q	$0.4268_{\pm 0.0589}$	$0.6763_{\pm 0.0718}$	$0.9197_{\pm 0.0924}$	$0.4931_{\pm 0.0762}$
0.1000	Aug	$0.5639_{\pm 0.1686}$	$1.1736_{\pm 0.5247}$	$1.5284_{\pm 0.5981}$	$0.9478_{\pm 0.3369}$
0.0300	Q	$0.4683_{\pm 0.0784}$	$0.7916_{\pm 0.1107}$	$1.0418_{\pm 0.1396}$	$0.5894_{\pm 0.1028}$
0.0300	Aug	$0.7291_{\pm 0.2718}$	$1.6897_{\pm 0.8463}$	$2.1367_{\pm 0.9472}$	$1.3726_{\pm 0.5834}$
0.0050	Q	$0.5497_{\pm 0.1163}$	$0.9748_{\pm 0.1719}$	$1.2389_{\pm 0.2047}$	$0.7436_{\pm 0.1537}$
0.0050	Aug	$1.0438_{\pm 0.4679}$	$2.7814_{\pm 1.4938}$	$3.4176_{\pm 1.6839}$	$2.2467_{\pm 1.0716}$