Generalized Mundlak Estimators with Network Interference

Gevorg Khandamiryan*

May 10, 2025

Abstract

Large language models, trained on personal data, may soon be able to mimic individual personalities. These "AI clones" or "AI agents" have the potential to transform how people search for matches in contexts ranging from marriage to employment. This paper presents a theoretical framework to study the tradeoff between the substantially expanded search capacity of AI clones and their imperfect representation of humans. An individual's personality is modeled as a point in k-dimensional Euclidean space, and an individual's AI clone is modeled as a noisy approximation of that personality. I compare two search regimes: In the "in-person regime," each person randomly meets some number of individuals and matches to the most compatible among them; in the "AI representation regime," individuals match to the person whose AI clone is most compatible with their AI clone. I show that a finite number of in-person encounters yields a better expected match than search over infinite AI clones. Moreover, when personality is sufficiently high-dimensional, simply meeting two people in person is more effective than search on an AI platform, regardless of the size of its candidate pool.

1. Introduction

The fixed effect model is a common approach to estimate causal parameters when there is group unconfoundedness. This makes sure that comparison of within group units is valid, and is reflected by a group-specific intercept in the equation below:

$$Y_i \;=\; \alpha_{g(i)} \;+\; D_i\,\tau \;+\; X_i^\top \beta \;+\; \varepsilon_i, \tag{1}$$

where D_i is the treatment assignment and $\alpha_{g(i)}$ is the fixed effect intercept of the group that unit i belongs to, $g(i)$. As extensively documented by Arkhangelsky and Imbens (2024), the credibility and

*UC Berkeley.

properties of the causal parameter τ depend on several factors and have certain caveats. Importantly, this specification rests on the additivity assumption of fixed effects and independent regressors, including the treatment.

Arkhangelsky and Imbens (2024) propose a Mundlak-style style interpretation of the fixed effects estimator, called Generalized Mundlak Estimators (GMEs). They suggest to think of the estimation procedure through a transformed expression of equation (1) that gets rid of the group-adjusting intercept and instead includes group-level averaged regressors in a following way:

$$Y_i \;=\; D_i\tau \;+\; \bar{D}_{g(i)}\gamma \;+\; X_i^\top\beta \;+\; \bar{X}_{g(i)}^\top\delta \;+\; \eta_i, \tag{2}$$

These enable several improvements over the fixed effect estimators in several aspects. First, they allow non-linear dependencies of group-averaged regressors to be incorporated. This framework enables us to create doubly robust estimators (Chernozhukov et al., 2017; Athey et al., 2018). And finally, one can include averages of compound functions of multiple regressors on a group level. The transition from equation (1) to equation (2) to achieve more robust estimators with a bigger capture of group-level heterogeneity comes at a cost of requiring a stronger statement than group unconfoundedness. It will now be sufficient to require the following:

$$W_i \;\perp\!\!\!\perp\; (Y_i(0), Y_i(1)) \mid X_i, \overline{H}_{g(i)},$$

where we adopt the potential outcomes framework of Neyman and Rubin (1974), $\overline{H}_{g(i)}$ is the balancing score. It is the average of a group-adjusting function $H(W_i, X_i)$, which ideally captures the unit differences between groups in a way that the above assumption holds. If we label the group that unit i belongs to as $L_{g(i)}$, then Arkhangelsky and Imbens (2024) argue that it is enough to reduce the unconfoundedness assumptions to one that is conditional on some balancing score that depends on regressors' group averages including non-linear relationships and mixed functions, rather than conditioning on the group label itself, such as $W_i \;\perp\!\!\!\perp\; (Y_i(0), Y_i(1)) \mid L_{g(i)}$. For example, one can structurally motivate that a sufficient condition is:

$$W_i \;\perp\!\!\!\perp\; (Y_i(0), Y_i(1)) \mid X_i, \overline{W}_{g(i)}, \overline{X}_{g(i)}, \overline{WX}_{g(i)}$$

If one decides to drop $\overline{WX}_{g(i)}$ from the conditioning set, then it means that they are confident that it is acceptable to compare treated and control units that have the same values of X_i and $\overline{S_{g(i)}} = \overline{W}_{g(i)}, \overline{X}_{g(i)}$, but which belong to groups that have different values of $\overline{WX}_{g(i)}$. Intuitively, conditioning on a balancing score allows us to compare groups with respect to some metric: using solely group labels would not allow one to to rank which groups are "closest" to each other. The average treatment estimator then is defined as:

$$\tau(s,x) \equiv \mathbb{E}[\tau_{L_{g(i)}}(X_i) \mid \overline{S}_{g(i)} = s,\, X_i = x] = \mathbb{E}[Y_i(1) - Y_i(0) \mid \overline{S}_{g(i)} = s,\, X_i = x].$$

This setup relies on the assumption of potential outcomes framework which excludes the presence of

spillovers across units and groups. In this paper I will tackle the problem of introducing interference in a setting where units are connected in a known network. This poses several issues that need to be addressed. First, in the usual setting the group membership of each unit i is known, and hence the construction of the balancing score relies on an averaging process where the groups are fixed. Upon introducing network interference, it is now an open problem as to how we should aggregate those balancing scores.

Since we condition on the balancing score to ensure some version of unconfoundedness, then the task of choosing the groups across which we will average the relevant confounders becomes a task of finding nodes in the network that are "similar" to each other according to some metric. This should have an intuitive interpretation in a way that one can easily make comparison statements such as "nodes in group 1 are closer to nodes in group 2 than to group 3". Once we know the relevant criteria according to which we will classify two or more nodes to be observationally equivalent, we will be able to define a deterministic group membership structure over the network. We will show that under a common network formation structure, a general labeling approach that uses the network adjacency matrix to establish a group membership protocol will be consistent with the Assumption 3.3 of Arkhangelsky and Imbens (2024) setup that requires an existence of sufficient statistic in a form of a balancing score, and will extend the implementation of GMEs to cases with network interference.

We will adopt inverse propensity weighting approach in a doubly robust setting: first, I will introduce a treatment vector which will incorporate both direct and indirect effects (this can be relaxed to a general exposure mapping setup), and then I will use the novel group labeling to construct a balancing score vector consisting of the averages of included regressors (including the treatment vector) with respect to the newly defined groups to construct propensity scores and use them in outcome regression framework for average treatment effect estimation. The latter now will have both direct and indirect treatment interpretation. The last section of this paper will implement a Monte Carlo simulation study to show the consistency of the proposed estimator.

2. Related literature

In essence, I propose pooling units with similar aggregate statistics similar in line with a general literature of grouped fixed-effects strategies (Hahn and Moon, 2010; Bonhomme and Manresa, 2015; Bonhomme et al., 2022). Also my approach is similar to the common panel data method of controlling for group heterogeneity (Mundlak, 1978, Chamberlain, 1984, Chamberlain, 2010, and Altonji and Matzkin, 2005). The recent approaches in fixed effects estimation (e.g. De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Wooldridge, 2021) circulate ideas that are related OLS estimators under heterogeneity by restricting conditional outcome distributions. Arkhangelsky and Imbens (2024) and this paper, however, leverage the treatment assignment structure to use cross-sectional approaches and adapt them to a group setting.

A relatively new and extensive literature on network data tackles problems of network formation

estimation (Graham 2016, 2017), network interference in causal settings (Leung, 2022, Forastiere 2017, 2022) and sampling (Chandrasekhar and Jackson, 2016). Graham (2017) proposes a homophily model of dyadic link formation with node-level heterogeneity. He then constructs a tetrad-logit estimator, which in essence, is an estimator for the homophily parameter that drives the link formation estimated from sampling tetrads from the overall network and seeing which one of those contribute to a score function based on their configuration. Cai (2023) studies linear regression models when the regressors include centrality measures of a sparse network and show that this can lead to inconsistent estimation procedures under measurement errors.

Forastiere et al. (2020) provide a generalized propensity score-based estimator to estimate treatment effects under network interference, but they do not adopt grouping strategy. They further investigate spillover effects on policy interventions in agricultural markets. Sanchez Becerra (2022) argues that a three dimensional general propensity score that consists of network adjacency matrix and node level covariates is enough to construct consistent doubly robust ATE estimators in networks with spillover concerns.

Leung and Loupos (2024) introduce a doubly robust estimation of ATE with network interference using GNNS in the first stage of constructing propensity scores. Leung (2022) introduces the notion of ANI (approximate neighborhood interference) which assumes that treatments assigned at further distances in the network have smaller but nonzero effects, contrary to the main approach in the literature. He shows that standard IPW estimators are consistent and approximately normal. Our approach of conditioning estimation procedure on network adjacency matrix is similar to Leung (2020) who derives restrictions on the network degree distribution under which spillover estimators are consistent and asymptotically normal verified with a strategic model of network formation.

Leung (2023) establishes a bias-variance trade-off results cluster-randomized trials with Cross-interference. He argues that constructing fewer, larger clusters reduces bias and increases variance, and gives a balancing rule for choosing the number of clusters from the adjacency matrix by simply looking at eigenvalues of it that are close to zero.

The use of some "similarity" criterion to control for heterogeneity across groups using group-level balancing scores is one of the key contributions in this paper. This approach has been popularized recently by Auerbach (2019, 2022), where he introduces the agent codegree function. This shows the probability that two units will have a common neighbor, and argues that conditioning on this function which serves as a similarity criterion, is sufficient to estimate parameters in partially linear models. This codegree function is consistently estimated by the squared adjacency matrix, which includes all of the identifiable information about the linking behavior in the majority of network formation models: hence, matching pairs of units with similar columns of this matrix enables consistent estimation of parameters which can include ATE estimators. Auerbach (2019) and Zeleneev (2020) focus on network pseudo-metrics to analyze general forms of unobserved heterogeneity in networks. Auerbach (2019) argues that this form of heterogeneity cannot be separately identified from spillovers in dense networks with exogenous treatment (Sanchez Becerra, 2022). This concern does not carry over to our setting, because most interesting networks are sparse

as far as interference is concerned in grouped data.

3. Model

I will propose a network formation model with logistic errors, which will be essential to the choice of balancing score. Assume N units connected in an undirected dyadic network. Let $A_{ij} = 1$ if i and j are connected and zero otherwise. Self-links are ruled out. Linking information in summarized in the $N \times N$ adjacency matrix $\mathcal{A}$, whose ij-th entry is A_{ij}. An agent's degree equals the number of links she has $A_{i+} = \sum_{j \neq i} A_{ij}$. $N \times 1$ vector $\mathcal{A}_+ = (A_{1+}, \ldots, A_{N+})'$, gives the network's degree sequence. Each unit i has an observable characteristic vector of agent-level attributes X_i. These are used to construct the $K \times 1$ pair vector $C_{ij} = g(X_i, Xj)$, which drives the link formation as shown next. Assume a network formation model:

$$A_{ij} \;=\; \mathbf{1}\Big(C_{ij}'\beta_0 \;+\; V_i \;+\; V_j \;-\; U_{ij} \;\geq 0\Big),$$

where V_i and V_j are the agent-level unobservable attributes and $U_{ij} = U_{ji}$ is the idiosyncratic component independently and identically distributed across pairs. This is in line with the transferable utility approach of link formation, i.e. all pairs with net positive utility form a link. The realization of error component U_{ij} fully and uniquely determines the network.

Assumption 1: Assume U_{ij} is a standard logistic random variable independently and identically distributed across pairs. Then *the conditional likelihood of observing the network* $\mathbf{A} = \mathbf{a}$ *is*

$$\Pr\big(\mathbf{A} = \mathbf{a} \mid \mathbf{X}, \mathbf{V}\big) \;=\; \prod_{i<j} \Pr\big(A_{ij} = a \mid X_i, X_j, V_i, V_j\big),$$

with

$$\Pr\big(A_{ij} = a \mid \mathbf{X}, \mathbf{V}\big) \;=\; \left[\frac{1}{1 + \exp\big(C_{ij}'\beta_0 + V_i + V_j\big)}\right]^{1-a} \left[\frac{\exp\big(C_{ij}'\beta_0 + V_i + V_j\big)}{1 + \exp\big(C_{ij}'\beta_0 + V_i + V_j\big)}\right]^{a}, \qquad \forall\, i \neq j.$$

Lemma 1: Under Assumption 1, the network degree sequence $\mathcal{A}_+ = (A_{1+}, \ldots, A_{N+})'$ is a sufficient statistic for unobservable agent-level unobservable attributes.

Proof: Following Graham (2017),

$$\Pr\big(\mathbf{A} = \mathbf{a} \mid \mathbf{X}, \mathbf{V}\big) \;=\; \prod_{i<j} \left[\frac{\exp\big(C_{ij}'\beta_0 + V_i + V_j\big)}{1 + \exp\big(C_{ij}'\beta_0 + V_i + V_j\big)}\right]^{a_{ij}} \left[\frac{1}{1 + \exp\big(C_{ij}'\beta_0 + V_i + V_j\big)}\right]^{1-a_{ij}} =$$

$$\textstyle\prod_{i<j} \frac{\exp\left[a_{ij}\left(C_{ij}'\beta_0 + V_i + V_j\right)\right]}{1+\exp\left(C_{ij}'\beta_0 + V_i + V_j\right)} \;=\; \underbrace{\Big\{\prod_{i<j}[1 + \exp\big(C_{ij}'\beta_0 + V_i + V_j\big)]^{-1}\Big\}}_{c(\mathbf{C};\beta_0,\mathbf{V})} \exp\Big(\textstyle\sum_{i<j} a_{ij} C_{ij}'\beta_0 + \sum_{i<j} a_{ij}(V_i +$$

$V_j)\Big) = c(\mathbf{C};\beta_0,\mathbf{V})\,\exp[\,S_1(\mathbf{a},\mathbf{C})'\beta_0]\,\exp[\,\mathcal{A}'_+\,\mathbf{V}]$, where $S_1(\mathbf{a},\mathbf{C}) \;=\; \sum_{i=1}^{N}\sum_{j<i} a_{ij}C_{ij}$, and $S_2(\mathbf{a}) \;=\; (a_{1+},\ldots,a_{N+})'$. □

Lemma 2: The conditional density $f(\mathbf{X},\mathbf{V} \mid \mathbf{a})$ follows exponential family distribution with sufficient statistic that is a function of network adjacency matrix and agent-level attributes.

Proof:

$$f(\mathbf{X},\mathbf{V} \mid \mathbf{a}) = \frac{f_{X,V}(\mathbf{X},\mathbf{V})\,Pr(\mathbf{A}=\mathbf{a} \mid \mathbf{X},\mathbf{V})}{\int f_{X,V}(x,v)\,Pr(\mathbf{A}=\mathbf{a} \mid x,v)\,dx\,dv}.$$

Given that

$$\Pr(\mathbf{A}=\mathbf{a} \mid \mathbf{X},\mathbf{V}) \;=\; c(\mathbf{C};\beta_0,\mathbf{V})\,\exp[\,S_1(\mathbf{a},\mathbf{C})'\beta_0]\,\exp[\,\mathcal{A}'_+\,\mathbf{V}],$$

we have:

$$f(\mathbf{X},\mathbf{V} \mid \mathbf{a}) \propto f_{X,V}(\mathbf{X},\mathbf{V})\,c(\mathbf{C};\beta_0,\mathbf{V})\,\exp\Big[S_1(\mathbf{a},\mathbf{C})^\top\beta_0 + \mathcal{A}_+^\top\mathbf{V}\Big].$$

Hence, $(S_1(\mathbf{a},\mathbf{X}),\ \mathcal{A}_+)$ is *sufficient* for $(\mathbf{X},\mathbf{V})$. □

Note that C and X can be used interchangeably. Intuitively, the previous results means that the network degree sequence is the best information that identifies heterogeneity. $S_1(\mathbf{a},\mathbf{X})$ is also a function of adjacency elements and agent-level attributes.

Now, suppose the regression to estimate the causal parameter with no interference is:

$$Y_i \;=\; \alpha_{g(i)} \;+\; D_i\,\tau \;+\; X_i^\top\beta \;+\; \varepsilon_i, \tag{3}$$

which modified to include GMEs becomes:

$$Y_i \;=\; D_i\tau \;+\; \bar{D}_{g(i)}\gamma \;+\; X_i^\top\beta \;+\; \bar{X}_{g(i)}^\top\delta \;+\; \eta_i \tag{4}$$

To account for interference and more general settings (such as linear-in-means model), define a new treatment vector and regressor vector as:

$$T_i \;=\; f(i,\,\mathbf{D},\,\mathbf{A}) \quad\text{and}\quad W_i \;=\; q(i,\,\mathbf{X},\,\mathbf{A}), \tag{5}$$

$$Y_i \;=\; g(T_i,\,W_i,\,\varepsilon_i) \quad\text{and}\quad \varepsilon_i \;\perp\!\!\!\perp\; T_i \mid W_i, \tag{6}$$

where T_i entirely summarizes interference while W_i summarizes confounding: $\mathbf{X}$ is the set of raw covariates, $\mathbf{A}$ is the adjacency matrix of the network, $\mathbf{D}$ is the treatment assignment. Common examples of T_i and W_i are

$$T_i \;=\; \Big(D_i,\,\sum_{j=1}^{n} A_{ij}D_j\Big) \quad\text{and}\quad W_i \;=\; \Big(X_i,\,\sum_{j=1}^{n} A_{ij},\,\frac{\sum_{j=1}^{n} A_{ij}X_j}{\sum_{j=1}^{n} A_{ij}}\Big).$$

In this example, the first element of T_i accounts for the direct effect while the second row can be

interpreted as the indirect effect of the treatment arising from having a treated neighbor. The confounding vector W_i can include both the raw X_i and its functions with the adjacency matrix. Note that you can think of the vector X here as the same agent-level attributes that drive link formation earlier. I left the notation identical to avoid confusion between the models. Hence, the main regression model can be formulated in terms of a treatment vector as follows:

$$Y_i \;=\; T_i\tau \;+\; \bar{T}_{g(i)}\gamma \;+\; W_i^{\top}\beta \;+\; \bar{W}_{g(i)}^{\top}\delta \;+\; \eta_i, \tag{7}$$

where $g(i)$ is the grouping variable and $S_{g(i)}$ is the group-level balancing score (can include averages of functions of T_i, W_i and so on). We can also combine them into one feature vector $(W_i,\; S_{g(i)})$.

Theorem 1: Suppose the grouped sampling assumption holds (see Arkhangelsky and Imbens (2024). Then, for the model defined in equations (4) - (6), such that each unit i in a large population is characterized by a pair of potential outcomes $(Y_i(0),\, Y_i(1))$ and a group label $L_{g(i)}$:

(i) $\{(T_i, W_i)\}_{i:g(i)=g} \;\perp\!\!\!\perp\; L_g \mid \bar{U}_g$,

(ii) $T_i \;\perp\!\!\!\perp\; \big(Y_i(0),\, Y_i(1)\big) \mid W_i,\; \bar{U}_{g(i)}$,

where $\bar{U}_g$ is an unobserved group-level variable that captures all the information about the distribution of (T_i, W_i): $\bar{U}_g \;\equiv\; \eta(L_g)$. Moreover, if Assumption 1 holds, then

$$T_i \;\perp\!\!\!\perp\; \big(Y_i(0),\, Y_i(1)\big) \mid W_i,\; \bar{S}_{g(i)},$$

where the balancing score $\bar{S}_{g(i)}$ is played by the role of the sufficient statistic of the conditional density of $f_{T_i,W_i|L_g}(t, w \mid l)$.

Proof: By Lemma 2, and given that $T_i \;=\; f(i,\, \mathbf{D},\, \mathbf{A})$ and $W_i \;=\; q(i,\, \mathbf{X},\, \mathbf{A})$, where $\mathbf{D}$ is the treatment that can potentially be endogenously determined by unobserved heterogeneity and the network structure, we have that $f(\mathbf{T}, \mathbf{W} \mid \mathbf{a}) \propto f(\mathbf{D}, \mathbf{X} \mid \mathbf{a})$, and has a sufficient statistic in the exponential family form comprising of $(S_1(\mathbf{a}, \mathbf{X}),\; \mathcal{A}_+)$, where $S_1(\mathbf{a}, \mathbf{C}) \;=\; \sum_{i=1}^{N}\sum_{j<i} a_{ij}C_{ij}$ and $\mathcal{A}_+ = (A_{1+}, \ldots, A_{N+})'$, is the network's degree sequence. Hence, by Lemma 1 in Arkhangelsky and Imbens (2024), we have that $\{(T_i, W_i)\}_{i:g(i)=g} \;\perp\!\!\!\perp\; L_g \,|\, \overline{U}_g$ and $T_i \;\perp\!\!\!\perp\; (Y_i(0), Y_i(1)) \,|\, W_i,\, \overline{U}_{g(i)},$, while Theorem 1 gives $T_i \;\perp\!\!\!\perp\; \big(Y_i(0),\, Y_i(1)\big) \mid W_i,\; \bar{S}_{g(i)}$. □

4. Estimation

In addition to the assumptions discussed in the previous section, the standard set of ATE assumptions are added.

Assumption 2: Grouped sampling: draw M groups, then $N_g \geq 2$ units per group.

Assumption 3: Known overlap: on a set A there is $\eta > 0$ with $\eta < e(W_i, S_{g(i)}) < 1 - \eta$.

Assumption 4: Moment bounds: $\mathbb{E}[\varepsilon_i(t)^2 \mid W_i, S_g] < K,\ \mathbb{E}[\varepsilon_i(t)^4] < \infty.$

Assumption 5: Rate conditions: cross-fitted estimators $\hat{\mu}, \hat{e}$ are mean-consistent and satisfy a product $o(M^{-1/2})$ rate.

The regular group-level unconfoundedness assumption $(T_i \perp (Y_i(0), Y_i(1)) \mid W_i, L_{g(i)})$ is now replaced with the exponential family assignment procedure, which is shown to be guaranteed by Assumption 1 that poses logistic errors on the link formation. The latter jointly with Theorem 1 assures that we can use a balancing score to ensure valid comparison of control and treated units, and also conjectures that the balancing score should be a function of the network degree sequence.

Under these assumptions, balancing on S_g identifies τ; the last two assumptions deliver $\sqrt{M}$-consistency and valid inference for the doubly-robust GME.

We first start by fitting any classifier to estimate the propensity score:

$$\hat{e}_i \;=\; \Pr\big(T_i = 1 \,\big|\, W_i,\ S_{g(i)}\big).$$

Then estimate the outcome regression $\hat{\mu}\big(t,\, W_i,\, S_{g(i)}\big)$ with AIPW framework:

$$\psi_i \;=\; \hat{\mu}\big(1,\, W_i,\, S_{g(i)}\big) \;-\; \hat{\mu}\big(0,\, W_i,\, S_{g(i)}\big) \;+\left(\frac{T_i}{\hat{e}_i} \;-\; \frac{1-T_i}{1-\hat{e}_i}\right)\Big(Y_i \;-\; \hat{\mu}\big(T_i,\, W_i,\, S_{g(i)}\big)\Big).$$

Computing group means gives the Generalized Mundlak estimator of the treatment effect with interference:

$$\bar{\psi}_g \;=\; \frac{1}{N_g} \sum_{i\,:\,g(i)=g} \psi_i, \qquad A \;=\; \frac{1}{M} \sum_{g=1}^{M} \frac{1}{N_g} \sum_{i\,:\,g(i)=g} \mathbf{1}(\hat{e}_i \in [\eta,\, 1-\eta])$$

$$\hat{\tau}_{\mathrm{GME}} \;=\; \frac{\displaystyle\sum_{g=1}^{M} \bar{\psi}_g}{A}.$$

5. Simulation Study

I simulate random graph according to Assumption 1. Agent-level attributes $\mathbf{X}$ are drawn uniformly from $U[-1, 1]$, unobserved heterogeneity terms are drawn from $V \sim \mathcal{N}(0, 1)$, U_{ij} idiosyncratic components are drawn from logistic distribution with mean 0 and variance 1. Treatment assignment is randomized but as a function of covariates and the network structure, even though simulations without this restriction also yield very similar results. Once the network is realized, I perform k-clustering where I will try values of $4, 6$ and 16 for the number of clusters. These are based on the network degree distribution, i.e. units with similar degrees will be in one cluster.

Once the clusters are determined, I will average out regressors and treatment vector for each group. I will use a treatment vector consisting of direct effect and an indirect effect, which is 1 if

at least one of the neighbors is treated. The outcomes will be generated according to true direct effect of 2 and an indirect effect of 1. The propensity scores will be constructed using logit fit, while the outcome regression using random forest, though results are also robust with regards to these specifications. Interestingly, as a robustness test I have tried to generate the network with normal errors U_{ij}, and these have resulted in much bigger biases. The results are presented for different number of units and replications with differing number of group clusters. Table 1 includes results on the direct effect, while Table 2 on indirect.

The key observation is that for a network with lower number of units, the number of clusters does not really matter, i.e. there is not enough heterogeneity in degrees of nodes. As the sample size goes up to 1000, there is insignificant but nonzero difference between results of running with different clusters. Overall, the estimator has achieves higher success for the direct effect, though indirect effects are also moderately well estimates. Even though in Figure 1, the distribution of estimates for the indirect effect are not normally distributed, this tendency diminshes as the number of simulations goes up.

		n = 400			n = 1000		
		$k = 4$	$k = 6$	$k = 16$	$k = 4$	$k = 6$	$k = 16$
$R = 1000$	Mean	2.004227	2.004227	2.004227	2.014146	2.014142	2.014142
	Bias	0.004227	0.004227	0.004227	0.014146	0.014142	0.014142
	Std Dev	0.106011	0.106011	0.106011	0.068009	0.068013	0.068013
	RMSE	1.009801	1.009801	1.009801	1.016421	1.016418	1.016418
$R = 200$	Mean	2.007625	2.007625	2.007625	2.011372	2.011364	2.011364
	Bias	0.007625	0.007625	0.007625	0.011372	0.011364	0.011364
	Std Dev	0.108428	0.108428	0.108428	0.073797	0.073816	0.073816
	RMSE	1.013413	1.013413	1.013413	1.014048	1.014041	1.014041

Table 1: Simulation results for direct effect by sample size (n), number of clusters (k) and number of replications (R).

		n = 400			n = 1000		
		$k = 4$	$k = 6$	$k = 16$	$k = 4$	$k = 6$	$k = 16$
$R = 1000$	mean	0.896339	0.896339	0.896339	1.092969	1.092843	1.092843
	bias	-0.103661	-0.103661	-0.103661	0.092969	0.092843	0.092843
	std	0.674063	0.674063	0.674063	0.543304	0.543347	0.543347
	rmse	0.781659	0.781659	0.781659	0.804050	0.803986	0.803986
$R = 200$	mean	0.889303	0.889303	0.889303	1.054082	1.054078	1.054078
	bias	-0.110697	-0.110697	-0.110697	0.054082	0.054078	0.054078
	std	0.696381	0.696381	0.696381	0.557581	0.557628	1.054078
	rmse	0.796290	0.796290	0.796290	0.785079	0.785110	0.785110

Table 2: Simulation results for indirect effect by sample size (n), number of clusters (k) and number of replications (R).

Table 3 presents results for a dense network. Here the minor difference between results is noticed between all three specifications. RMSE, bias and standard error all go up in this case, consistent with Auerbach (2019) critic that exogenous treatment in dense networks poses risk of identifying heterogeneity separately from spillovers.

		Direct			Indirect		
		$k = 4$	$k = 6$	$k = 16$	$k = 4$	$k = 6$	$k = 16$
$R = 200$	mean	2.390844	2.383705	2.352578	0.393227	0.384366	0.355153
	bias	0.390844	0.383705	0.352578	-1.606773	-1.615634	-1.644847
	std	0.396173	2.383705	0.372892	0.383777	0.377957	0.357569
	rmse	1.445896	1.437746	1.402791	0.397428	0.394345	0.384964

Table 3: Simulation results for dense networks, direct and indirect effects reported. Sample size is 1000.

Figure 1 shows the distribution of estimates for direct and indirect effects in multiple specifications. Overall, indirect effect estimates do poorly in all specification, with better results as the sample size of nodes or the number of replications goes up.

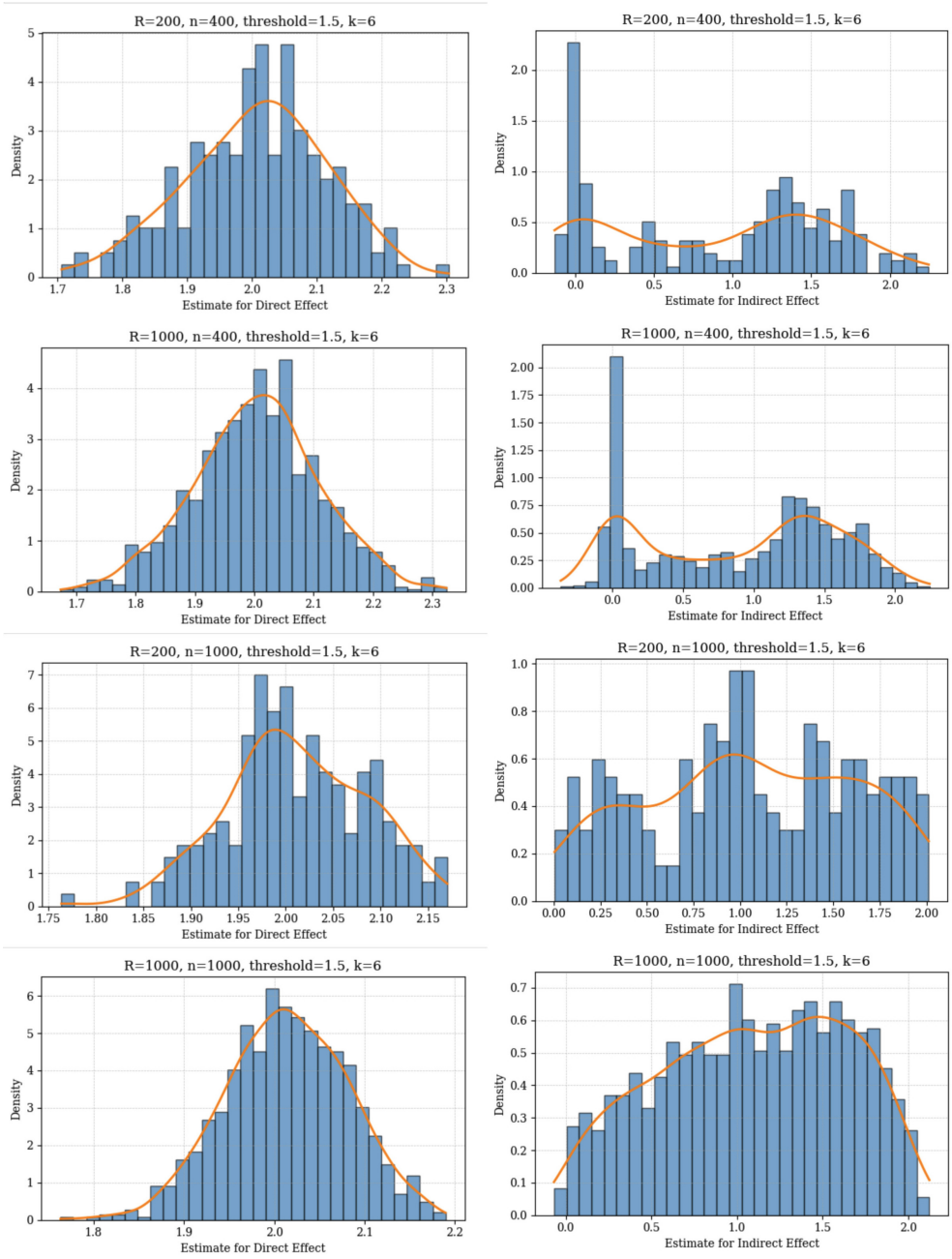

Figure 1: Density plots for simulation results.

6. Discussion and Conclusion

This paper proposed an approach to account for network interference in the Generalized Mundlak framework. The identification argument relies on the exponential family assumption from Arkhangelsky and Imbens (2024) and a new assumption on the link formation process in the network, a logistic error assumption.

Dividing units into clusters based on a "similarity" criteria which uses the network adjacency

matrix, we proposed an AIPW procedure with a balancing score that is a group-level average of included regressors. By leveraging a convenient sufficient statistic that results from logistic link formation models, we could justify group membership according to node degree. Monte Carlo experiments have confirmed the consistency and asymptotic behavior of this estimator.

References

[1] Altonji, J. G. and Matzkin, R. (2005). *Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors*. Econometrica.

[2] Arkhangelsky, D. and Imbens, G. (2024). *Fixed Effects and the Generalized Mundlak Estimator*. Review of Economic Studies.

[3] Athey, S., Imbens, G., and Wager, S. (2018). *Approximate residual balancing: debiased inference of average treatment effects in high dimensions*. Journal of the Royal Statistical Society.

[4] Auerbach, S. (2022). *Identification and Estimation of a Partially Linear Regression Model using Network Data*. Econometrica.

[5] Bonhomme, S. and Manresa, E. (2015). *Grouped Patterns of Heterogeneity in Panel Data*. Econometrica.

[6] Bonhomme, S., Lamadon, T., and Manresa, E. (2022). *Discretizing Unobserved Heterogeneity*. Econometrica.

[7] Cai, Y. (2023). *Linear Regression with Centrality Measures*. JMP.

[8] Chamberlain, G. (1984). *Panel Data*. Handbook of Econometrics.

[9] Chamberlain, G. (2010). *Binary Response Models for Panel Data: Identification and Information*. Econometrica.

[10] Chandrasekhar, A. G. and Jackson, M. O. (2016). *Econometrics of Sampled Networks*

[11] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2017). *Double/debiased machine learning for treatment and structural parameters*. The Econometrics Journal.

[12] De Chaisemartin, C. and d'Haultfoeuille, X. (2020). *Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects*. American Economic Review.

[13] Forastiere, L. (2017). *Identification and Estimation of Treatment and Interference Effects in Observational Studies on Networks*. Journal of the American Statistical Association.

[14] Forastiere, L. (2022). *Estimating Causal Effects under Network Interference with Bayesian Generalized Propensity Scores*. Journal of Machine Learning Research.

[15] Forastiere, L., Airoldi, E., and Mealli, F. (2020). *Identification and Estimation of Treatment and Interference Effects in Observational Studies on Networks*. Journal of the American Statistical Association.

[16] Goodman-Bacon, A. (2021). *Difference-in-Differences with Variation in Treatment Timing*. Journal of Econometrics.

[17] Graham, B. (2016). *Homophily and Transitivity in Dynamic Network Formation*. NBER Working Paper.

[18] Graham, B. (2017). *An Econometric Model of Network Formation with Degree Heterogeneity*. Econometrica.

[19] Hahn, J. and Moon, H. R. (2010). *Panel Data Models with Finite Number of Multiple Equilibria*. Econometrics Theory.

[20] Leung, M. P. (2020). *Treatment and Spillover Effects Under Network Interference*. Review of Economics and Statistics.

[21] Leung, M. P. (2022). *Causal Inference under Approximate Neighborhood Interference*. Econometrica.

[22] Leung, M. P. (2023). *Network Cluster-Robust Inference*. Econometrica.

[23] Leung, M. P. and Loupos, C. (2024). *Graph Neural Networks for Causal Inference Under Network Confounding*. Working Paper.

[24] Mundlak, Y. (1978). *On the Pooling of Time Series and Cross-Section Data*. Econometrica.

[25] Neyman, J. and Rubin, D. B. (1974). *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*. Journal of Educational Psychology.

[26] Sanchez Becerra, S. (2022). *The Network Propensity Score: Spillovers, Homophily, and Selection into Treatment*. JMP.

[27] Wooldridge, J. M. (2021). *Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators*. Working Paper.

[28] Zeleneev, A. (2020). *Identification and Estimation of Network Models with Nonparametric Unobserved Heterogeneity*. JMP.