Political Analysis (2023), **0**, 1–18
doi:10.1017/pan.2023.40

ARTICLE

Inference in Linear Dyadic Data Models with Network Spillovers

Nathan Canen[1,2,3] and Ko Sugiura[1]

[1]Department of Economics, University of Houston, Science Building, 3581 Cullen Boulevard Suite 230, Houston, TX 77204-5019, USA; [2]Department of Economics, University of Warwick, Coventry CV4 7AL, UK; [3]Research Economist, National Bureau of Economic Research, Cambridge, MA, USA

Corresponding author: Nathan Canen; Email: ncanen@uh.edu

(Received 9 February 2023; revised 18 May 2023; accepted 27 June 2023)

Abstract

When using dyadic data (i.e., data indexed by pairs of units), researchers typically assume a linear model, estimate it using Ordinary Least Squares, and conduct inference using "dyadic-robust" variance estimators. The latter assumes that dyads are uncorrelated if they do not share a common unit (e.g., if the same individual is not present in both pairs of data). We show that this assumption does not hold in many empirical applications because indirect links may exist due to network connections, generating correlated outcomes. Hence, "dyadic-robust" estimators can be biased in such situations. We develop a consistent variance estimator for such contexts by leveraging results in network statistics. Our estimator has good finite-sample properties in simulations, while allowing for decay in spillover effects. We illustrate our message with an application to politicians' voting behavior when they are seating neighbors in the European Parliament.

Keywords: dyadic data; networks; inference; cross-sectional dependence; congressional voting

Edited by: Jeff Gill

1. Introduction

Dyadic data are categorized by the dependence between two sets of sampled units (dyads). For example, exports between the United States and Canada depend on both countries (and, plausibly, their characteristics). This contrasts to classical data in the social sciences that only depend on a single unit of observation (e.g., the GDP of the United States, or a politician's vote in a roll call).

The empirical relevance of dyadic data is showcased by its widespread use, which has increased over the past two decades (Graham (2020a) provides an extensive review). For example, applications are found in political economy (correlation in voting behavior in Parliament across seating neighbors; Harmon, Fisman, and Kamenica 2019), international political economy and trade (export–import outcomes across countries; Anderson and van Wincoop 2003), and international relations (Hoff and Ward 2004; for a salient example), among many others. In fact, dyadic data are considered to be dominant in quantitative international relations (Poast 2016). In these examples, applied researchers typically model the dependence between dyadic outcomes and observable characteristics using a linear model, which they then estimate using Ordinary Least Squares (OLS). However, inference on such estimators for the linear parameters is more complex.

The main approach in recent applied work has been the use of the so-called "dyadic-robust" estimators (e.g., Aronow, Samii, and Assenova 2015; Cameron *et al.* (2011); Tabord-Meehan 2019, among others). Such estimators build on the widely used assumption in dyadic data that the error terms

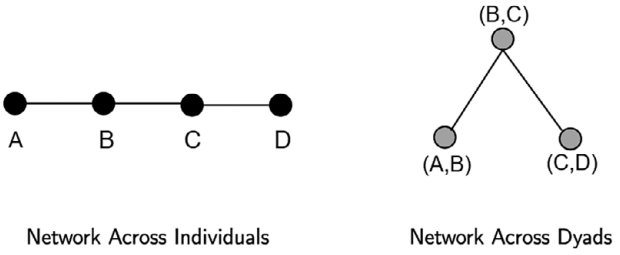

Figure 1. Hypothetical example of a network in Parliament.
Note: The left figure shows a hypothetical example of politician networks based on seating arrangements: *A* sits beside *B*, who sits beside *C*, who sits beside *D*. The right-hand figure illustrates the resulting network among active dyads. As dyads (A, B) and (B, C) share a unit, they are indirectly linked in the dyadic network. However, though dyads (A, B) and (C, D) do not have a politician in common, they might still be correlated through two indirect links: namely, *B* sits beside *C*, who sits beside *D*. Hence, *D*'s actions can affect politician *A*.

for dyad (i,j) and for dyad (k,l) can only be correlated if they share a unit (see Aronow *et al.* 2015; Tabord-Meehan 2019, for a discussion; Cameron and Miller 2014, for a review).

In this paper, we first argue that such an assumption does not hold in many applications using dyadic data where dyads may be indirectly connected along a network.[1] Figure 1 presents a simple example in the context of politicians in Congress, whose votes or decisions depend on their seating neighbors. It is completely possible that behavior across dyads (A,B) and (C,D) might be correlated along unobservables because they have many indirect connections (in the figure, through *A* sitting next to *B*, who sits next to *C*).[2] We show that such spillovers invalidate the assumptions for consistency of existing "dyadic-robust" variance estimators through generating interdependence, implying that they are biased for the true asymptotic variance when dyads may be correlated even when they do not share a common unit.

To deal with these issues, we develop a consistent variance estimator that explicitly accounts for such network spillovers even with dyadic data, thereby complementing existing approaches (e.g., Aronow *et al.* 2015).[3] We prove that our proposed variance estimator is consistent for the true variance of the OLS estimator in linear models with dyadic data when the cross-sectional dependence follows an observed (exogenous) network. Our main insight is that the dependence across all dyads, including indirect spillovers, can be rewritten as correlations across a specific network over dyads. This allows us to apply the framework of Kojevnikov, Marmer, and Song (2021) to such network random variables, although here it is a network over dyads, rather than individuals. Monte Carlo simulations show that our proposed estimator has good finite-sample properties and outperforms other estimators for the relevant contexts.

To help practitioners, we then provide a step-by-step guideline on whether our estimator may be appropriate to their context. As we describe, this choice depends on (i) whether spillovers from indirectly connected dyads are likely to be present, (ii) whether the researcher observes/constructs the network among dyads through which spillovers propagate, and (iii) whether those spillovers are likely to be persistent. Our variance estimator is consistent for the asymptotic variance of the OLS estimator

[1]This is a concrete class of applied examples where the assumption fails. The possibility that cross-sectional dependence in dyadic data might be more extensive than assumed has been pointed out by Cameron and Miller (2014) and Cranmer and Desmarais (2016).

[2]We expand on these examples in the next section and in the Supplementary Material. Such spillovers could be further rationalized as individual-level unobserved heterogeneity: for example, an unmeasured preference for voting Yes, or a preference for trading with a certain country (see Graham 2020b and references therein for details).

[3]We provide an extensive comparison of the relative benefits of each approach in the next section. We note here, though, that neither approach subsumes the other, as they depend on different assumptions and may be more appropriate for different applications.

even under (i)–(iii). And, our estimator can account for decay in propagation, as Corollary 3.1 and Example 3.1 illustrate.

Finally, we illustrate the extent to which neglecting network spillovers with dyadic data may bias inference results. Beyond Monte Carlo simulations, we revisit the application in Harmon *et al.* (2019) of voting in the European Parliament.[4] The authors study whether random seating arrangements (based on naming conventions) induce neighboring politicians to agree with one another in policy votes. The outcome, whether politicians i and j vote the same way on a policy, is dyadic in nature. However, i and j's votes may be positively correlated even if they are not neighbors: for instance, i and j may sit on either side of common neighbors k and l, who influence them both, and this seating arrangement is observed. This chain of influences is sufficient to induce strong positive correlation across non-dyads. We show that neglecting such higher-order spillovers has significant empirical consequences: their estimated variance using the estimator in Aronow *et al.* (2015) is roughly 22% smaller than using our consistent estimator accounting for such spillovers, while the estimate based on the Eicker–Huber–White estimator ignoring spillovers is approximately 73% smaller than our proposal, consistent with the arguments of Erikson, Pinto, and Rader (2014).

1.1. Related Literature

The use of dyadic data in Political Science has a rich history, particularly in International Relations. However, empirical challenges with such models are well known (see Poast 2016 for a historical overview). Early on, the concerns were mostly about model specification, including the error term. This includes the 2001 special issue of *International Organization*, mostly focusing on the use of fixed effects. More recently, Erikson *et al.* (2014) pointed out that ignoring dependence across dyads can lead to erroneous hypothesis testing, as computed standard errors would be too small. Hoff and Ward (2004), Minhas, Hoff, and Ward (2019), and Minhas *et al.* (2022) suggest including random coefficients and latent variables to account for dependencies across dyads. Our approach explicitly accounts for the whole network of interdependencies across dyads, which can go beyond third-order dependences (assumed in Minhas *et al.* 2019, 2022). It does so by using asymptotic inference, rather than Bayesian (Minhas *et al.* 2019, 2022) or randomized inference (Erikson *et al.* 2014).

As a result, our paper is directly related to the literature on (asymptotic) inference in regression analysis with dyadic random variables. Aronow *et al.* (2015) and Tabord-Meehan (2019) consider OLS estimation and inference in a linear dyadic regression model. Meanwhile, Graham (2020a) and Graham (b) explore a likelihood-based approach to dyadic regression models, while Graham, Niu, and Powell (2022) and Chiang and Tan (2023) provide results for kernel density estimation in dyadic regression models. It is also related to other developments in multiway clustering, as we detail in Appendix A.2 of the Supplementary Material. While useful to allow for correlations along time and within such groups, this separable structure may be inappropriate for environments where spillovers follow a complex form of dependence along a network.

However, we emphasize that neither approach subsumes the other. The papers cited above leave the dependence within "clusters" (groups of dyads that share units) unrestricted, but assume independence across such clusters. This is akin to the literature with one-way clustering (e.g., Hansen and Lee 2019). By comparison, our approach restricts such dependence among groups of dyads that share units (i.e., dependence is assumed to follow the observed network), but allows for dependence across such "clusters" of dyads along the dyadic network.

2. Setup

Assume that we observe a cross section of $N \in \mathbb{N}$ individuals located along a network—the latter interpretable as politicians, countries, firms, or other observation units depending on the context. The

[4] Replication materials for all results are available online in Canen and Sugiura (2023).

dyads present in the N-individual network (i.e., among the $\binom{N}{N-1}$ possible dyads) are called active dyads, so that the dyad for two units i and j (e.g., politicians and countries) is denoted by some m. The set of active dyads is denoted by $\mathcal{M}_N$ and M denotes the cardinality of that set.

We assume that each dyad m is endowed with a triplet of dyad-specific variables, forming a triangular array $\{(y_{M,m}, x_{M,m}, \varepsilon_{M,m})\}_{m \in \mathcal{M}_N}$ with respect to M, where $y_{M,m} \in \mathbb{R}$ is a one-dimensional observable outcome, $x_{M,m} \in \mathbb{R}^K$ is a K-dimensional vector of observable characteristics with $K \in \mathbb{N}$, and $\varepsilon_{M,m} \in \mathbb{R}$ is a one-dimensional random error term that is not observable to the researcher. We only consider exogenous network formation and the network is assumed to be observable. These conditions are summarized in the following assumption.

Assumption 2.1 (Exogenous and Observable Dyadic Networks). The network among dyads is assumed to be conditionally independent of $\{\varepsilon_{M,m}\}_{m \in \mathcal{M}_N}$. Furthermore, this network among the N individuals is assumed to be observable.

While such assumptions are standard in models of dyadic networks, they seem particularly appropriate when units or dyad pairs are linked across geographical, physical, or ex ante social relations (e.g., family ties). This includes capturing neighboring and regional spillovers across countries, as often done in international relations, or exogenous seating arrangements in Parliament, as illustrated in the examples in the next section.

The subsequent arguments require us to distinguish between a pair of dyads who share a member (i.e., who are directly linked—which we call, *adjacent*) and a pair of dyads who are directly or indirectly linked (which we call, simply, *connected*).

Definition 1 (Adjacent and Connected Dyads). Two active dyads m and m' are said to be adjacent if they have an individual in common; and they are called connected if they are linked through pairs of adjacent dyads.

In Figure 1, dyad (A,B) is adjacent to (B,C), and connected with, though not adjacent to, (C,D). Hence, the adjacency relationship constitutes a network structure among active dyads, and thus a network over individuals can be transformed to one over active dyads. For example, the right-hand side panel of Figure 1 provides a network over pairs of voting politicians (i.e., active dyads).[5] We define the geodesic distance between two connected dyads m and m' to be the smallest number of adjacent dyads between them. Note that adjacent dyads are a special case of connected dyads with geodesic distance equal to one.

2.1. The Linear Model

2.1.1. Setup and Identification

The cross-sectional model of interest takes the form of the linear network-regression model: for any $N \in \mathbb{N}$,

$$y_{M,m} = x'_{M,m}\beta + \varepsilon_{M,m} \quad \forall m \in \mathcal{M}_N, \tag{1}$$

where

$$Cov(\varepsilon_{M,m}, \varepsilon_{M,m'} \mid X_M) = 0 \quad \text{unless } m \text{ and } m' \text{ are connected}, \tag{2}$$

and β is a $K \times 1$ vector of the regression coefficients and X_M denotes the $M \times K$ matrix that records the observed dyad-specific characteristics, that is, $X_M \coloneqq [x_{M,1}, \ldots, x_{M,M}]'$.

[5] This corresponds to thinking about the line graph of the original graph over individuals.

In this paper, we assume that β is identified, which follows from standard assumptions on strict exogeneity, lack of multicollinearity and the existence of finite second moments of $y_{M,m}$ and $x_{M,m}$. (For completeness, see Assumption B.1 and Proposition B.1 in the Supplementary Material.)

We note that equation (2) allows for there to be spillovers across the error terms even when dyads m and m' are not adjacent, as long as they are connected through indirect links. By comparison, applied researchers such as Harmon *et al.* (2019) and Lustig and Richmond (2020) (and the estimators of Aronow *et al.* (2015) and Tabord-Meehan (2019)) consider a variant of the linear regression (1) under the assumption

$$Cov(\varepsilon_{M,d(i,j)},\varepsilon_{M,d(k,l)} \mid x_{M,d(i,j)},x_{M,d(k,l)}) = 0 \quad \text{unless} \quad \{i,j\} \cap \{k,l\} \neq \varnothing, \tag{3}$$

with $m = d(i,j)$ representing the dyad between i and j. This specific assumption would be equivalent to setting

$$Cov(\varepsilon_{M,m},\varepsilon_{M,m'} \mid X_M) = 0 \quad \text{unless } m \text{ and } m' \text{ are adjacent.} \tag{4}$$

2.1.2. Examples

Whether to allow indirect spillovers (as in (2)) or not (as in (4)) depends on the researchers' applications. We now present examples where our approach may be preferable.

Example 2.1 (Gravity Model of Bilateral Trade Flow). A researcher is studying the trade flow from country i to j, with (log) exports from i to j denoted y_{ij}. Following the literature, (s)he assumes y_{ij} follows the structural gravity equation (e.g., Anderson and van Wincoop 2003; Eaton and Kortum 2002; Helpman, Melitz, and Rubinstein 2008; Melitz 2003):

$$y_{ij} = \alpha + \beta z_{ij} + \gamma \sum_{k \neq i} g_{ki} y_{ki} + \eta_{ij}, \tag{5}$$

where z_{ij} represents a dyadic characteristic of i and j, such as the shipping cost, whether both countries are democratic (e.g., Mansfield, Milner, and Rosendorff 2000), or whether both participate in WTO/GATT (e.g., Gowa and Kim 2005); $\sum_k g_{ki} y_{ki}$ is the amount i spends on imports (g_{ki} equals one if country i purchases goods from country k and zero otherwise), and η_{ij} captures unobserved heterogeneity pertaining to the trade flow between countries i and j.

To see our main point, suppose there are only four countries (1, 2, 3, and 4) which trade, where country 1 exports to country 2, which in turn exports to country 3, and country 3 exports to country 4. Equation (5) then simplifies to $y_{12} = \alpha + \beta z_{12} + \eta_{12}$, $y_{23} = \alpha + \beta z_{23} + \gamma y_{12} + \eta_{23}$, and henceforth.

Rearranging these equations implies that the trade flow from country 3 to country 4 can be written as

$$y_{34} = \alpha + \alpha\gamma + \alpha\gamma^2 + \gamma^2\beta z_{12} + \gamma\beta z_{23} + \beta z_{34} + \gamma^2\eta_{12} + \gamma\eta_{23} + \eta_{34}.$$

Therefore, $Cov(y_{12},y_{34} \mid z) = \gamma^2 Var(\eta_{12} \mid z) \neq 0$, where $z \equiv \{z_{12},z_{23},z_{34}\}$. Hence, there can be nonzero correlation between trade flows y_{12} and y_{34} even if they do not have a country in common. This is because an idiosyncratic shock to an upstream country can propagate through the trade network.

Example 2.2 (Legislative Voting). A researcher is interested in whether seating arrangements in legislatures can affect a politician's behavior, y_i (e.g., propensity to vote "Yes" on a roll call, as Harmon *et al.* 2019, or the amount of co-sponsoring, as Lowe and Jo 2021; Saia 2018, among others). For concreteness, suppose there are four politicians with the seating arrangements given by Figure 1.

The researchers posit that i's behavior can be influenced by the (average) of its seating neighbors' own voting behavior through a parameter γ as follows:

$$y_A = \alpha + \gamma y_B + \eta_A, \qquad y_B = \alpha + \gamma \frac{y_A + y_C}{2} + \eta_B, \tag{6}$$

$$y_C = \alpha + \gamma \frac{y_B + y_D}{2} + \eta_C, \qquad y_D = \alpha + \gamma y_C + \eta_D. \tag{7}$$

If $\gamma \neq 0$, A is affected by their neighbor B, while B is affected by both of its neighbors (A and C) and so forth. The researcher is interested in whether neighbors' decisions are more highly correlated than the decisions among non-neighbors.

Denote y_{ij} as the dyadic outcome of interest (e.g., a measure of correlation between i and j's decisions). Both y_{AB} and y_{CD} involve y_B and y_C, which are themselves a function of η_B and η_C. Hence, $Cov(y_{AB}, y_{CD}) \neq 0$, even if the two pairs of legislators do not share a common member.

2.1.3. Estimation

Throughout this paper, we focus on the OLS estimator of β, denoted by $\hat{\beta}$. Under the assumptions above, we can write

$$\hat{\beta} - \beta = \left(\sum_{j \in \mathcal{M}_N} x_{M,j} x'_{M,j} \right)^{-1} \sum_{m \in \mathcal{M}_N} x_{M,m} \varepsilon_{M,m}. \tag{8}$$

It is straightforward to verify that $\hat{\beta}$ is unbiased for β under our identification conditions (Assumption B.1 in the Supplementary Material). However, a consistency result is by no means trivial due to the dependence along the network which induces a complex form of cross-sectional dependence, hindering a naïve application of the standard theory for independently and identically distributed (*i.i.d.*) random vectors.

2.2. Outline of Our Procedure

2.2.1. Inference

Inference about β is based on a normal approximation of the distribution of $\hat{\beta}$ around β. We focus on hypothesis testing conducted using the expression

$$\left(\widehat{Var}(\hat{\beta}) \right)^{-\frac{1}{2}} (\hat{\beta} - \beta), \tag{9}$$

where $\widehat{Var}(\hat{\beta})$ is a consistent estimator of the asymptotic variance of $\hat{\beta}$. Our main result in Section 3.4 is providing such an appropriate estimator, which takes the form

$$\widehat{Var}(\hat{\beta}) := \left(\sum_{k \in \mathcal{M}_N} x_k x'_k \right)^{-1} \left(\sum_{m \in \mathcal{M}_N} \sum_{m' \in \mathcal{M}_N} \kappa_{m,m'} h_{m,m'} \hat{\varepsilon}_m \hat{\varepsilon}_{m'} x_m x'_{m'} \right) \left(\sum_{k \in \mathcal{M}_N} x_k x'_k \right)^{-1}, \tag{10}$$

where $\kappa_{m,m'}$ is an appropriate kernel function that will formally be defined in Section 3.4; $h_{m,m'}$ represents an indicator function that takes one if dyads m and m' are connected and zero otherwise; and $\hat{\varepsilon}_m := y_m - x'_m \hat{\beta}$.

This paper derives conditions under which $\widehat{Var}(\hat{\beta})$ is consistent for the asymptotic variance of $\hat{\beta}$. Before doing so, let us compare the variance estimator (10) with an often used estimator based on one-way clustering of dyad groupings.

Remark 2.1 (Dyadic-Robust Variance Estimator). An increasing number of applied researchers, such as Harmon *et al.* (2019) and Lustig and Richmond (2020), estimate model (1) and conduct inference using the following dyadic-robust variance estimators proposed by Aronow *et al.* (2015) and Tabord-Meehan (2019):

$$\widehat{Var}(\hat{\beta}) := \left(\sum_{k \in \mathcal{M}_N} x_k x'_k \right)^{-1} \left(\sum_{m \in \mathcal{M}_N} \sum_{m' \in \mathcal{M}_N} \mathbb{1}_{m,m'} \hat{\varepsilon}_m \hat{\varepsilon}_{m'} x_m x'_{m'} \right) \left(\sum_{k \in \mathcal{M}_N} x_k x'_k \right)^{-1}, \tag{11}$$

where $\mathbb{1}_{m,m'}$ equals one if dyads m and m' are adjacent and zero otherwise.

Note that the use of the dyadic-robust variance estimator sets cases in which two dyads are not adjacent, but connected, to zero. Meanwhile, our estimator (10) accounts for network spillovers by accommodating the correlation across both adjacent and connected dyads.[6] As our examples above suggest, the structure of the variance estimator (11) may not be compatible with indirect spillovers in some settings, which should assume the specification (2) instead. This suggests that the dyadic-robust variance estimator may be inconsistent when non-adjacent dyads can still affect the correlation structure and outcomes of dyad m.[7] This conjecture is formally proven in Corollary 3.1 and illustrated in Monte Carlo simulations in Section 4. We note that this is a feature of applying such dyadic-robust variance estimators to network spillovers, and not a feature of those estimators per se.

2.2.2. Guidelines on Whether and How to Use the Proposed Estimator

1. When deciding whether to use our proposed estimator (10), the researcher should first ask whether spillovers from indirectly connected dyads are likely to be present (and not decay immediately) in their setup: that is, is equation (2) a more appropriate assumption than equation (4)?

   While this depends on the specific application, Examples 2.1 and 2.2 illustrate models where that is likely to be the case. And, condition (17) provides a notion of how much persistence is needed for a bias to appear. As we show below, these insights are robust to decaying spillover effects (see Corollary 3.1, Example 3.1, and the associated simulation results).
2. If such spillovers of unobservables are likely to exist, are they governed by an exogenous and observable network (Assumption 2.1), such as physical, geographical, or social (e.g., family ties)? If so, the proposed estimator is appropriate under regularity conditions.
3. One can implement our estimator by (i) choosing a kernel (e.g., rectangular; see Section 3.4), (ii) setting the lag-truncation, b_M (either by a known value, or adaptively by $b_M = 2\log(M)/\log(\max(\textit{average degree}, 1.05))$, where M is the number of dyads and we use the average degree of the dyadic network, and (iii) plugging-in those choices into equation (10).

   The estimator is consistent under regularity conditions, even when spillovers decay, and shows good finite-sample properties in the simulations below.

3. Theoretical Results

3.1. Network-Dependent Processes

Let $Y_{M,m}$ be a random vector defined as

$$Y_{M,m} \coloneqq x_{M,m}\varepsilon_{M,m},^{8} \tag{12}$$

and denote $\mathcal{C}_M \coloneqq \{x_{M,m}\}_{m\in\mathcal{M}_N}$.[9] From equations (8) and (12), we can write

$$\hat{\beta} - \beta = \left(\frac{1}{M}\sum_{j\in\mathcal{M}_N} x_{M,j}x'_{M,j}\right)^{-1}\frac{1}{M}\sum_{m\in\mathcal{M}_N} Y_{M,m}. \tag{13}$$

[6] See Definition 1 and the subsequent discussion. The choice of kernel and lag-truncation is discussed in Section 3.4.

[7] Clustering estimators may be inappropriate when the correlation structure has network spillovers as in (2), because each agent has a complex (i.e., non-separable) structure of connections, reflected in a non-separable network across dyads. If the network model features positive spillovers, then the dyadic-robust variance estimator will likely underestimate the true variance, leading to conservative hypothesis testing. Meanwhile, it is likely to overstate the true variance when there are negative spillovers. We expand on this point in our numerical simulations.

[8] By construction, the collection of $Y_{M,m}$'s constitutes a triangular array of random vectors.

[9] For the case of stochastic networks, it is defined to include information about the network topology as well as the collection of the dyad-specific attributes $\{x_{M,m}\}_{m\in\mathcal{M}_N}$. See Kojevnikov *et al.* (2021).

Our interest lies in proving the asymptotic properties of $\hat{\beta}$ taking advantage of the expression (13). However, the presence of $\varepsilon_{M,m}$ in $Y_{M,m}$, which is allowed to be correlated along the network over active dyads, renders our approach nonstandard and unsuitable for applications of canonical results, such as those for *i.i.d.* random variables or even other variants, including spatially correlated and time-series data.

The main insight of this paper is that the spillovers across connected—even if not adjacent—dyads can be rewritten as the dependence of $Y_{M,m}$'s along the network of active dyads (hereby, referred to as the "network"). This allows us to embrace such complex cross-sectional dependence and appropriately rewrite the problem so that recent results on network dependent random variables (Kojevnikov *et al.* 2021) can be applied. To do so, the dependence between random variables for any two sets of dyads A and B, $Y_{M,A}$ and $Y_{M,B}$, which are at a distance s from one another, is assumed to be controlled by a sequence of bounded (random) coefficients $\theta_{M,s}$. As the minimum distance, s, between A and B grows, the dependence $\{\theta_{M,s}\}$ between $Y_{M,A}$ and $Y_{M,B}$, is assumed to go to zero. A formal description is provided in Appendix A of the Supplementary Material.

3.2. Definitions

As will become transparent shortly, asymptotic theories for $\hat{\beta}$ rest on tradeoffs between the correlation of the network-dependent random vectors (i.e., the dependence coefficients) and the denseness of the underlying network. To measure the denseness, we first define two concepts of neighborhoods: for each $m \in \mathcal{M}_N$ and $s \in \mathbb{N} \cup \{0\}$,

$$\mathcal{M}_N(m;s) \coloneqq \{m' \in \mathcal{M}_N : \rho_M(m,m') \leq s\},$$
$$\mathcal{M}_N^{\partial}(m;s) \coloneqq \{m' \in \mathcal{M}_N : \rho_M(m,m') = s\},$$

where $\rho_M(m,m')$ denotes the geodesic distance between dyads m and m'.[10] The former set collects all the m's neighbors whose distance from m is no more than s (which we call a neighborhood), while the latter registers all the m's neighbors whose distance from m is exactly s (which we call a neighborhood shell).

Next, we define two types of density measures of a network: for $k, r > 0$,

$$\begin{aligned}\Delta_M(s,r;k) &\coloneqq \frac{1}{M}\sum_{m\in\mathcal{M}_N}\max_{m'\in\mathcal{M}_N^{\partial}(m;s)}|\mathcal{M}_N(m;r)\backslash\mathcal{M}_N(m';s-1)|^k,\\ \delta_M^{\partial}(s;k) &\coloneqq \frac{1}{M}\sum_{m\in\mathcal{M}_N}|\mathcal{M}_N^{\partial}(m;s)|^k,\end{aligned} \tag{14}$$

where it is assumed that $\mathcal{M}_N(m';-1) = \varnothing$. The former measure gauges the denseness of a network in terms of the average size of a version of the neighborhood. Kojevnikov *et al.* (2021) show that controlling the asymptotic behavior of an appropriate composite of these two measures (denoted by c_M and defined in Assumption 3.6) is sufficient for the Law of Large Numbers (LLN) and Central Limit Theorem (CLT) of the network dependent random variables (Condition ND).

Note that there are two different units at play here: the number of sampling units (i.e., individuals), N, and the number of dyads, M. We now assume that $M \to \infty$ as $N \to \infty$, eliminating the possibility of extremely sparse networks among individuals. This is empirically relevant and consistent with both applied and theoretical literatures (see Appendix A.1 of the Supplementary Material for a discussion).

Assumption 3.1. $M \to \infty$ as $N \to \infty$.

[10] Recall that we define the geodesic distance between two connected dyads m and m' to be the smallest number of adjacent dyads between them.

3.3. Asymptotic Properties of $\hat{\beta}$

We make use of the following two regularity assumptions for the proof of consistency of $\hat{\beta}$ for β (Theorem 3.1) and to derive its asymptotic distribution (Theorem 3.2).[11] All proofs can be found in Appendix B of the Supplementary Material.

Assumption 3.2 (Conditional Finite Moment of ε_m). There exists $p > 4$ such that $\sup_{N\geq 1} \max_{m\in\mathcal{M}_N} E\left[|\varepsilon_m|^p \mid \mathcal{C}_M\right] < \infty$ a.s.

Assumption 3.3 (Kojevnikov *et al.* 2021, Assumption 3.4). There exists a positive sequence $r_M \to \infty$ such that for $k = 1, 2$,

$$\frac{M^2\theta_{M,r_M}^{1-1/p}}{\sigma_M} \xrightarrow{a.s.} 0, \quad \text{and} \quad \frac{M}{\sigma_M^{2+k}} \sum_{s\geq 0} c_n(s, r_M; k)\theta_{M,s}^{1-\frac{2+k}{p}} \xrightarrow{a.s.} 0,$$

as $M \to \infty$, where $p > 4$ is the same as in Assumption 3.2.

Assumption 3.2 requires that the errors are not too large once conditioned on common shocks. Together with the standard full-rank assumption for identification of β, this implies Assumption 3.1 of Kojevnikov *et al.* (2021) for each uth element of $Y_{M,m}$, denoted by $Y_{M,m}^u$ with $u \in \{1, \ldots, K\}$.

Assumption 3.3 is a condition that controls the tradeoff between the denseness of the underlying network and the covariability of the random vectors. If the network becomes dense, then the dependence of the associated random variables has to decay much faster. This embodies the idea that spillovers decay as they propagate farther (see, e.g., Kelejian and Prucha (2010)), which is consistent with the applications described above. For instance, Acemoglu, García-Jimeno, and Robinson (2015) assume that network spillovers are zero if agents are sufficiently distantly connected on a geographical network. This assumption may be violated for very dense networks with low decay of spillovers.

3.3.1. Consistency

Theorem 3.1 (Consistency of $\hat{\beta}$). *Under Assumptions 3.1–3.3,* $||\hat{\beta} - \beta||_2 \xrightarrow{p} 0$ *as* $N \to \infty$.

When Assumption 3.1 is dropped, Theorem 3.1 continues to hold in terms of the number of active dyads M.

3.3.2. Asymptotic Normality

Let $S_M \coloneqq \sum_{m\in\mathcal{M}_N} Y_{M,m}$, which is present in $\hat{\beta}$ in equation (13). Let S_M^u be the uth entry of S_M for $u \in \{1, \ldots, K\}$ and denote the unconditional variance of S_M^u by $\tau_M^2 \coloneqq Var(S_M^u)$. Since S_M^u is not a sum of independent variables, its variance cannot be simply expressed as a sum of the variances of $Y_{M,m}$. We thus need to explicitly take into account covariance between the random variables $\{Y_{M,m}^u\}_{m\in\mathcal{M}_N}$. We study the CLT for the normalized sum of $Y_{M,m}^u$, which is given by $\frac{S_M^u}{\tau_M}$.

Assumption 3.4 bridges the conditional variance (assumed in Assumptions 3.2 and 3.3) and the unconditional variance of $\frac{S_M}{\tau_M}$, which we are interested in. The final assumption for the asymptotic normality result is a standard regularity condition guaranteeing that the asymptotic variance is well defined,[12] which follows from both matrices in the expression being well-defined.

Assumption 3.4 (Growth Rates of Variances). $\frac{\sigma_M^2}{\tau_M^2} \xrightarrow{a.s.} 1$ as $N \to \infty$.

[11] These assumptions are required for Theorem 3.2, but as usual, the proof of consistency (Theorem 3.1) can be derived under weaker conditions. (See Assumptions B.2 and B.3 and their associated discussion, in the Supplementary Material.)

[12] Further note that Theorem 3.2 is proved under a weaker condition than Assumption 3.4.

Assumption 3.5. (a) For all $N \geq 1$, $\{x_{M,m}\}_{m \in \mathcal{M}_N}$ have uniformly bounded support.
(b) $\lim_{N\to\infty} \left(\frac{1}{M} \sum_{k \in \mathcal{M}} E\left[x_{M,k} x'_{M,k}\right] \right)$ is positive definite.
(c) $\lim_{N\to\infty} \frac{N}{M^2} \sum_{m \in \mathcal{M}_N} \sum_{m' \in \mathcal{M}_N} E\left[\varepsilon_{M,m} \varepsilon_{M,m'} x_{M,m} x'_{M,m'}\right]$ exists with finite elements.

Under these assumptions, the asymptotic distribution of $\hat{\beta}$ is given by the following.

Theorem 3.2 (Asymptotic Normality of $\hat{\beta}$). *Under Assumptions 3.1–3.5,* $\sqrt{N}\left(\hat{\beta} - \beta\right) \xrightarrow{d} \mathcal{N}(0, AVar(\hat{\beta}))$ *as* $N \to \infty$, *where*

$$AVar(\hat{\beta}) = \lim_{N\to\infty} N \left(\sum_{k \in \mathcal{M}_N} E\left[x_{M,k} x'_{M,k}\right] \right)^{-1} \left(\sum_{m \in \mathcal{M}_N} \sum_{m' \in \mathcal{M}_N} E\left[\varepsilon_{M,m} \varepsilon_{M,m'} x_{M,m} x'_{M,m'}\right] \right) \left(\sum_{k \in \mathcal{M}_N} E\left[x_{M,k} x'_{M,k}\right] \right)^{-1}, \tag{15}$$

which is positive semidefinite with finite elements.

3.4. Consistent Estimation of the Asymptotic Variance of $\hat{\beta}$ under Network Spillovers

Our objective is to consistently estimate $AVar(\hat{\beta})$ defined in Theorem 3.2. As errors are mean zero, $Y_{M,m}$ is centered, that is, $E\left[Y_{M,m}\right] = 0$ for each $m \in \mathcal{M}_N$.

3.4.1. The Estimator

The proposed estimator is a type of kernel estimator. Let b_M denote the bandwidth, or the lag truncation (its choice is described in Section 3.4.2) and $\omega : \overline{\mathbb{R}} \to [-1,1]$ a kernel function such that $\omega(0) = 1$, $\omega(z) = 0$ whenever $|z| > 1$, and $\omega(z) = \omega(-z)$ for all $z \in \overline{\mathbb{R}}$. The feasible variance estimator of interest is

$$\widehat{Var}(\hat{\beta}) = \left(\frac{1}{M} \sum_{k \in \mathcal{M}_N} x_{M,k} x'_{M,k} \right)^{-1} \left(\frac{1}{M^2} \sum_{s \geq 0} \sum_{m \in \mathcal{M}_N} \sum_{m' \in \mathcal{M}_N^{\partial}(m;s)} \omega_M(s) \hat{Y}_{M,m} \hat{Y}'_{M,m'} \right) \left(\frac{1}{M} \sum_{k \in \mathcal{M}_N} x_{M,k} x'_{M,k} \right)^{-1}, \tag{16}$$

with $\omega_M(s) \coloneqq \omega\left(\frac{s}{b_M}\right)$ for all $s \geq 0$ and $\hat{Y}_{M,m} \coloneqq x_{M,m} \hat{\varepsilon}_{M,m}$, where $\hat{\varepsilon}_{M,m} \coloneqq y_{M,m} - x'_{M,m} \hat{\beta}$.

3.4.2. Choice of Lag Truncation, b_M

There are two approaches for the choice of the associated lag truncation parameter. First, the researcher may already know (or is willing to impose) the truncation, perhaps due to a theoretical/institutional motivation. For instance, Acemoglu *et al.* (2015) set the lag to two in a related problem. Then, the thought exercise is that this choice will adapt as $M \to \infty$ according to the assumptions below. Alternatively, the researcher could use a data-driven choice. Assumption 3.6(c) suggests that it should depend on both the sample size and the network topology, including the average degree of the dyadic network. One such selection rule is suggested in Kojevnikov *et al.* (2021) based on their proofs: $b_M = 2\log(M)/\log(\max(\textit{average degree}, 1.05))$.

3.5. Consistency of the Proposed Estimator

The consistency of the variance estimator requires two sets of additional assumptions. The first set is Assumption 4.1 of Kojevnikov *et al.* (2021), but stated here in terms of the network over dyads.

Assumption 3.6 (Kojevnikov *et al.* 2021, Assumption 4.1). There exists $p > 4$ such that

(a) $\sup_{N\geq 1} \max_{m\in\mathcal{M}_N} E\left[|\varepsilon_m|^p \mid \mathcal{C}_M\right] < \infty a.s.$;

(b) $\lim_{M\to\infty} \sum_{s\geq 1} |\omega_M(s) - 1|\delta_M^{\partial}(s)\theta_{M,s}^{1-\frac{2}{p}} = 0 a.s.$;

(c) $\lim_{M\to\infty} \frac{1}{M} \sum_{s\geq 0} c_M(s, b_M; 2)\theta_{M,s}^{1-\frac{4}{p}} = 0 a.s.$, where

$$c_M(s,r;k) \coloneqq \inf_{\alpha>1} \left(\Delta_M(s,r;k\alpha)\right)^{\frac{1}{\alpha}} \left(\delta_M^{\partial}\left(s; \frac{\alpha}{\alpha-1}\right)\right)^{\frac{\alpha-1}{\alpha}}.$$

Assumption 3.6(a) is a stronger counterpart to Assumption 3.2, as it requires that a higher-order (i.e., higher than fourth order) conditional moment be well-defined. Assumption (b) posits a tradeoff between the kernel function, the denseness of a network, and the dependence coefficients. Specifically, the kernel function ω_M is required to converge to one sufficiently fast. Kojevnikov *et al.* (2021) demonstrate primitive conditions under which this requirement is fulfilled (Proposition 4.2). Assumption (c) requires that the correlation coefficients decay much faster relative to the denseness of the network. This is satisfied in the suggested choice for b_M above.

Another set of conditions restricts the denseness of the network, ruling out the situation where the network becomes progressively dense: most notably, the case where every single individual unit is directly linked to every other individual.

Assumption 3.7. (a) $\sup_{N\geq 1} \sum_{s\geq 0} \delta_M^{\partial}(s;1) < \infty$; (b) $\lim_{M\to\infty} \frac{1}{M} \sum_{s\geq 0} c_M(s, b_M; 2) = 0$.

The following theorem is the main theoretical contribution of this paper.

Theorem 3.3 (Consistency of the Network-Robust Variance Estimator)**.** *Under the conditions for Theorem 3.2, and Assumptions 3.6 and 3.7,* $\left\|N\widehat{Var}(\hat{\beta}) - Var(\hat{\beta})\right\|_F \xrightarrow{p} 0$ *as* $N \to \infty$*, where* $\|\cdot\|_F$ *indicates the Frobenius norm.*

Theorem 3.3 establishes the consistency of our proposed variance estimator accounting for network spillovers across dyads in the sense of the Frobenius norm.

3.6. When to Use the Proposed Estimator and the Role of Decaying Spillover Effects?

It follows from Theorem 3.3 that the dyadic-robust variance estimator (11) is inconsistent for the true variance when the underlying network involves a non-negligible degree of far-away correlations, as suggested in the examples of the previous section.[13] Specifically, the following corollary states that the dyadic-robust variance estimator of Aronow *et al.* (2015) may not necessarily be consistent when it is naïvely applied to the network-regression model with nonzero correlations beyond direct neighbors.

Corollary 3.1 (Inconsistency of Dyadic-Robust Estimators with Network Spillovers)**.** *Suppose that the assumptions required in Theorem 3.3 hold. Assume, in addition, that*

$$\inf_{N\geq 1} \frac{1}{M} \left\| \sum_{s\geq 2} \sum_{m\in\mathcal{M}_N} \sum_{m'\in\mathcal{M}_N^{\partial}(m;s)} E\left[\varepsilon_{M,m}\varepsilon_{M,m'}x_{M,m}x'_{M,m'}\right] \right\|_F > 0. \tag{17}$$

Then, the dyadic-robust estimator (11) applied to the network-regression models (1) and (2) is inconsistent.

[13] If the network is such that there are only adjacent dyads (i.e., when equation (4) holds), then the result above implies consistency of this estimator for dyadic dependence. By comparison, Lemma 1 of Aronow *et al.* (2015) and Propositions 3.1 and 3.2 of Tabord-Meehan (2019) also provide consistent variance estimators for the dyadic dependence case without higher-order network spillovers. However, these results and ours do not subsume one another. Indeed, their estimators can accommodate flexible dependence within clusters of dyads that share common units, while we assume that the network of spillovers is observed even if there are only adjacent connections.

The added condition (17) in Corollary 3.1 pertains to both the network configuration of active dyads and the regression variables. It represents a setting where the spillovers from far-away neighbors are non-negligible even when N is large. For instance, (17) can hold even if there are not many neighbors, as long as the covariances between the error terms are sufficiently large. This builds on Erikson *et al.* (2014) that inference with dyadic data may be biased if one only partially accounts for such spillovers. On the other hand, if far-away neighbors in the network have only a negligible effect on the cross-sectional dependence, the dyadic-robust variance estimator remains a good approximation for the asymptotic variance of linear dyadic data models with network spillovers across dyads. These insights are investigated further in Section 4 using numerical simulations.

These observations extend to settings where the spillovers decay along the network (i.e., when the correlation along unobservables decreases with the geodesic distance among dyads). Indeed, our estimator already accounts for such decay through the indirect covariances in its expression (16). When such spillovers propagate and decay is not too high, then condition (17) is satisfied, as such spillovers are non-negligible. On the other hand, if they decay at a very high rate (in the limit, a 100% decay from adjacent to connected dyads), then our estimator will become very similar to the dyadic-robust variance estimator.

However, some researchers may be willing to tolerate some asymptotic bias to still implement the dyadic-robust estimator. Then, when should they prefer our proposed estimator? While a general answer is complex because the bias depends on both the strength of indirect spillovers and the network configuration, the example below, together with the subsequent simulations, provides useful directions for salient settings.

Example 3.1 (Maximum Admissible Bias in the Dyadic-Robust Variance Estimator). Suppose that spillovers decay exponentially with distance along the network: i.e., $E\left[\varepsilon_{M,m}\varepsilon_{M,m'}x_{M,m}x_{M,m'}\right] = \gamma^s$, where s is the geodesic distance between dyads m and m', $\gamma \in (0,1)$ and S the longest path in the network.

Let $B > 0$ denote the maximum tolerance for condition (17) that the researcher is willing to allow when using the dyadic-robust variance estimator (11). Then, a sufficient condition for the researcher to prefer the proposed estimator (16) over (11) is that the decay rate γ is higher than a threshold $\bar{\gamma}$, where

$$\ln \bar{\gamma} = \frac{2}{S+2}\left\{\ln B - \ln(S-1) - \frac{1}{S-1}\sum_{s\geq 2}\ln \delta_M^{\partial}(s)\right\}.$$

When the tolerated bias is small ($B \to 0$), dependence is large enough and does not decay too fast (large γ), or the network is more dense, our approach is preferable because it provides a consistent estimator even with non-negligible spillovers. Since the network is observed, S and the last term are estimable and can be used for such diagnostics. Appendix C.5 of the Supplementary Material provides further discussions, while the next section presents for our baseline results and discusses simulation exercises.

4. Monte Carlo Simulations

4.1. Simulation Design

We compare three types of variance estimators across different specifications and network configurations. We use the Eicker–Huber–White estimator as a benchmark,[14] the dyadic-robust estimator of Tabord-Meehan (2019) as a comparison accounting for the dyadic nature of the data (when inappropriately used in the presence of network spillovers), and our proposed estimator which is robust to network spillovers across dyads.

We first generate networks on which random variables are assigned by following Canen, Schwartz, and Song (2020), among others, by employing two models of random graph formations. They are

[14] It is used in Bliss and Russett (1998) and Mansfield *et al.* (2000), for instance.

referred to as Specifications 1 and 2. Specification 1 uses the Barabási and Albert (1999) model of preferential attachment, with the fixed number of edges $\nu \in \{1,2,3\}$ being established by each new node.[15] Specification 2 is based on the Erdös–Renyi random graph (Erdös and Rényi 1959, 1960) with probability $p = \frac{\lambda}{N}$ for N denoting the number of nodes and $\lambda \in \{1,2,3\}$ being a parameter that governs the probability relative to the node size. The summary statistics for the networks generated by Specifications 1 and 2 are given in Appendix C.1 of the Supplementary Material.

For each of the randomly generated networks, the simulation data are generated from the following simple network-linear regression:

$$y_m = x_m\beta + \varepsilon_m,^{16}$$

with $m \coloneqq d(i,j)$ representing the dyad between agent i and j. The dyad-specific regressor x_m is defined as $x_m \coloneqq |z_i - z_j|$, where both z_i and z_j are drawn independently from $\mathcal{N}(0,1)$. The regression coefficient is fixed to $\beta = 1$ across specifications.

The dyad-specific error term ε_m is constructed to exhibit nonzero correlation with $\varepsilon_{m'}$ as long as dyads m and m' are connected (i.e., in the network terminology, there exists a path in the simulated network), while the strength of the correlation is assumed to decay as they are more distant. This decay is parametrized by γ (see Appendix C of the Supplementary Material for details). If $\gamma = 1$, then spillover effects are the same no matter how far the agents are apart, that is, the spillover effects do not decay. If $\gamma = 0$, there are no spillover effects, so the dyadic-robust variance estimator should be consistent. The case of $S = 2$ corresponds to a situation where up to friends of friends may matter for spillovers.

We consider three scenarios for each type of network. In the main text, we set $S = 2$ and $\gamma = 0.8$. The results for $S = 2$ with $\gamma = 0.2$ are given in Appendix C.5 of the Supplementary Material, and the ones for $S = 1$ with $\gamma = 0.8$ are in Appendix C.6 of the Supplementary Material. For comparison purposes, we employ the mean-shifted (by one) rectangular kernel with the lag truncation equal to two throughout the experiments.

4.2. Results

In Table 1, we present the coverage probability for β and the average length of the confidence interval across simulations. To do so, we compute the t-statistic using the OLS estimator for β and different variance estimators under a normal distribution approximation.[17] The finite-sample properties of the three variance estimators are further illustrated in Figure 2 in Appendix C.3 of the Supplementary Material.

The results for the empirical coverage probabilities depend on two dimensions: the sample size (N) and the denseness of the underlying network (parametrized by ν and λ). The coverage probability for each estimator improves with the sample size. However, when spillovers are high ($\gamma = 0.8$), only our proposed network-robust variance estimator has coverage close to 95%, consistent with Theorem 3.3. Meanwhile, in this setup, both the Eicker–Huber–White and the dyadic-robust variance estimators perform poorly as the underlying network becomes denser, no matter which specification of the network is involved. For example, in Specification 1 with $\nu = 3$ and the largest sample size ($N = 5{,}000$), the confidence intervals based on the Eicker–Huber–White and the dyadic-robust variance estimators do not cover the true parameter 615 and 455 times out of 5,000 simulations (12.3% and 9.1%),

[15] In generating the Barabási–Albert random graphs, we follow Canen *et al.* (2020) by choosing the seed to be the Erdös–Renyi random graph with the number of nodes equal the smallest integer above $5\sqrt{N}$, where N denotes the number of nodes.

[16] To simplify notation, we drop the M subscript, making the triangular array structure implicit.

[17] It is well known that the estimates of a variance–covariance matrix may be negative semidefinite when the sample size is very small. This occurs in 4 out of 5,000 simulations when $N = 500$. Rather than dropping such observations, we follow Cameron *et al.* (2011) and augment the eigenvalues of the matrix by adding a small constant, say 0.005, thereby obtaining a new variance estimate that is more conservative.

Table 1. The empirical coverage probability and average length of confidence intervals for β at 95% nominal level: $S = 2$, $\gamma = 0.8$.

		Specification 1			Specification 2		
	N	$\nu = 1$	$\nu = 2$	$\nu = 3$	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$
		Coverage probability					
Eicker–Huber–White	500	0.877	0.868	0.871	0.891	0.870	0.875
	1,000	0.880	0.873	0.873	0.892	0.881	0.888
	5000	0.879	0.871	0.877	0.890	0.882	0.880
Dyadic-robust	500	0.922	0.898	0.894	0.932	0.921	0.917
	1,000	0.929	0.913	0.901	0.937	0.927	0.924
	5,000	0.934	0.912	0.909	0.939	0.933	0.922
Network-robust	500	0.930	0.917	0.915	0.937	0.937	0.941
	1,000	0.939	0.934	0.933	0.946	0.945	0.948
	5,000	0.949	0.944	0.943	0.947	0.948	0.948
		Average length of the confidence intervals					
Eicker–Huber–White	500	0.368	0.409	0.482	0.287	0.285	0.296
	1,000	0.266	0.302	0.331	0.205	0.201	0.207
	5,000	0.132	0.159	0.176	0.092	0.090	0.094
Dyadic-robust	500	0.426	0.454	0.520	0.328	0.329	0.337
	1,000	0.312	0.339	0.361	0.236	0.232	0.237
	5,000	0.158	0.178	0.192	0.106	0.104	0.108
Network-robust	500	0.441	0.493	0.568	0.337	0.349	0.366
	1,000	0.326	0.373	0.408	0.244	0.248	0.259
	5,000	0.167	0.199	0.222	0.110	0.112	0.118

Note: The upper-half of the table displays the empirical coverage probability of the asymptotic confidence interval for β, and the lower-half showcases the average length of the estimated confidence intervals. As the sample size (N) increases, the empirical coverage probability for our estimator accounting for network spillovers approaches 0.95, the correct nominal level. However, that is not the case for alternative estimators.

respectively. On the other hand, the network-robust variance estimator is designed to capture higher-order correlations and, thus, its coverage remains stable across network configurations.

A similar conclusion is drawn from the average length of the confidence intervals: the confidence intervals for the Eicker–Huber–White and dyadic-robust variance estimators are typically 10%–20% shorter than those for our proposed estimator when γ is large and $S = 2$. This means the former undercovers the true parameter (in the presence of positive spillovers). However, as the magnitude of spillovers decreases (i.e., γ tends to zero), higher-order spillovers are less pronounced, so that the biases from using the Eicker–Huber–White and dyadic-robust variance estimators disappear. This is shown in Table 7 of Appendix C.5 of the Supplementary Material for the case of $S = 2$ and $\gamma = 0.2$. When $S = 1$, the dyadic-robust variance estimator coincides with our proposed estimator (i.e., there are no spillovers from non-adjacent links, or spillovers fully decay immediately). This is shown in Table 8 of Appendix C.6 of the Supplementary Material.

Finally, Appendix C.7 of the Supplementary Material shows that the results are robust to spillovers that can reach the most distantly connected neighbors ($S = \infty$) and to choosing the lag-truncation adaptively, following the rule $b_M = 2\log(M)/\log(\max(\textit{average degree}, 1.05))$ suggested above.

5. Empirical Illustration: Legislative Voting in the European Parliament

We now turn to an empirical application, revisiting the work of Harmon *et al.* (2019) on whether legislators who sit next to each other in Parliament tend to vote more alike on policy proposals.

They focus on the European Parliament, whose Members (MEPs) are voted in through elections in each European Union (EU) member country every 5 years. The Parliament convenes once or twice a month, in either Brussels or Strasbourg, to debate and vote on a series of proposals. Once elected to the European Parliament (EP), these MEPs are organized into European Political Groups (EPGs), which aggregate similar ideological members/parties across countries. As Harmon *et al.* (2019) describe, these EPGs function as parties for many of the traditional party functions in other legislatures, including coordination on policy and policy votes. Most importantly, MEPs sit within their EPG groups in the chamber. However, within each EPG group, non-party leaders traditionally sit in alphabetical order by last name. See Figure 4 in Appendix D.1 of the Supplementary Material for an example.

5.1. Data

We adopt the dataset used in Harmon *et al.* (2019), which collects the MEP-level data on votes cast in the EP. The dataset records what each MEP voted for (Yes or No), where she was seated, and a number of individual characteristics (e.g., country, age, education, gender, and tenure). We restrict the sample to the policies voted in Strasbourg during the seventh term, and we focus on the seating pattern between July 14 and July 16, 2009 (which involved 116 different proposals being voted on). The resulting sample has 2,431,261 observations, which are split into 422 politicians forming 26,099 pairs (i.e., dyads) of MEPs over 116 proposals.[18] Further information on the construction of our sample is detailed in Appendix D.3 of the Supplementary Material.

5.2. Empirical Setup

We follow Harmon *et al.* (2019) in assuming that two MEPs who are seated next to each other within the same political group are treated as an active dyad and that such relations are exogenously determined. Their main specification is a linear model:

$$Agree_{d(i,j),t} = \beta_0 + \beta_1 SeatNeighbors_{d(i,j),t} + \varepsilon_{d(i,j),t}, \tag{18}$$

where $Agree_{d(i,j),t}$ is an indicator that takes one if MEP i and j cast the same vote on proposal t, and zero otherwise, $SeatNeighbors_{d(i,j),t}$ is a binary variable that equals one if MEP i and j are seated next to each other when the vote for proposal t is taken place, and zero otherwise. The authors originally conducted inference using the estimator in Aronow *et al.* (2015), assuming that dyads $m = d(i,j)$ cannot be correlated with $m' = d(k,l)$ unless they share a common member.

We compare this approach to using the variance estimator introduced in Section 3.4, which allows the error terms to exhibit nonzero correlations as long as they are connected on the network over dyads represented by the adjacency relation of seating arrangements in Parliament. We use the mean-shifted rectangular kernel with the lag truncation equal the longest path in the constructed network, which accommodates all the possible correlations across connected dyads (i.e., pairs of MEPs), placing equal weight on each of them.[19]

Inspired by Harmon *et al.* (2019), we consider three specifications: (I) a simple linear regression model as given in (18); (II) the model (18) augmented with a flexible set of other demographic variables;[20] and (III) the model (18) with both a flexible set of other demographic variables and

[18] There are 334 pairs of adjacent dyads and 591 pairs of connected dyads.

[19] In Appendix D.4 of the Supplementary Material, we replicate this analysis with a different choice of kernel and setting the lag-truncation parameter following the criterion suggested above/in Kojevnikov *et al.* (2021). The results are very similar.

[20] Following Harmon *et al.* (2019), we include indicators whether country of origins, quality of education, freshman status, and gender, respectively, are the same, as well as differences in ages and tenures.

Table 2. Spillovers in legislative voting—main analysis

	Specification (I)	Specification (II)	Specification (III)
Panel A: Parameter estimates			
seat neighbors	0.007	0.006	0.006
Panel B: Standard errors			
Eicker–Huber–White	0.003	0.003	0.003
Dyadic-robust	0.008	0.008	0.009
Network-robust	0.010	0.010	0.011

Note: Panel A displays the parameter estimates for the variable "Seat Neighbors" for the three different specifications, and Panel B shows the standard errors for its regression coefficient using different variance estimators. Adjacency of MEPs is defined at the level of a row-by-EP-by-EPG. (See the note below Figure 4 in Appendix D.1 of the Supplementary Material.) The independent variables are Seat neighbors, whether both MEPs are from the same country; whether both MEPs have the same quality of education, whether both MEPs are freshman or not; the difference in the MEPs' ages; and the difference in the MEPs' tenures. A full description of the result is provided in Supplementary Table 13.

day-specific fixed effects. When fixed effects are present in their original estimation, we estimate a within-difference model via OLS.

5.3. Results

The main results of our empirical analysis are summarized in Table 2. Panel A displays the parameter estimates for the three different specifications. This panel shows that our point-estimates are consistent with the original estimates of Harmon *et al.* (2019) (columns 6 and 7 of Table 4), as they are close to 0.006 (their original results) and stable across specifications.[21] Hence, changes to point estimates are not due to sample selection. The positive coefficient for SeatNeighbors suggests that the MEPs sitting together tend to vote more similarly than those sitting apart, providing evidence in favor of their original hypothesis. The coefficients on the covariates (displayed in Panel C of Supplementary Table 13) are also quantitatively and qualitatively similar to those in their original paper.

Panel B shows the standard errors for the regression coefficient of SeatNeighbors using different variance estimators. As foreshadowed in the Monte Carlo simulations, the Eicker–Huber–White estimates are the smallest, followed by the dyadic-robust estimates, which, in turn, are smaller than the network-robust estimates. In fact, for Specification (III), the Eicker–Huber–White estimate is roughly 73% smaller than using the estimator accounting for network spillovers across dyads, while the dyadic-robust one is 22% smaller. This fact entails two implications. First, our finding provides empirical evidence in support of the existence of *indirect* positive spillovers among the MEPs: even distant connections may indirectly generate correlated behavior among politicians i and j. Second, the use of alternative estimators not accounting for such spillovers undercovers the true parameter and may generate biased hypothesis testing about the regression coefficient of SeatNeighbors. The difference in estimates appears quantitatively meaningful in this empirical example.

6. Conclusion

To conclude, we clarify that our goal in this exercise is neither to criticize dyadic-robust variance estimators, which are a fundamental part of the empiricist's toolkit, nor to suggest that our approach

[21] Note that our dependent variable is equal to one if two MEPs vote the same and zero otherwise, while Harmon *et al.* (2019) code it as one if MEPs vote differently. Hence, to compare our estimates with theirs, the signs on the estimates of *SeatNeighbors* must be flipped.

should always be used. Rather, we wish to draw attention that researchers should fully specify the cross-sectional dependence in their model. If the conventional assumption of dyadic dependence correctly specifies the environment in question, or when spillovers beyond immediate neighbors might be negligible, then previous approaches suffice. However, as we have discussed above, existing applications may apply the latter method even if it is seemingly inappropriate to their setting. This includes situations where such network spillovers may be present or persistent (even with decay). In such scenarios, our estimator provides a possible solution. Those choices, though, must be guided by the application that empiricists face. Hence, building on Poast (2016), we recommend researchers to continue to fully specify their model, including full specification of their covariance structure, thereby clarifying what type of inference procedure is most appropriate for their environment.

Acknowledgments. We thank Aimee Chin, Hugo Jales, Taisuke Otsu, Pablo Pinto, Vítor Possebom, Kevin Song, Bent E. Sorensen, and seminar participants at the University of Houston, the 2022 Texas Econometrics Camp, the 2022 European Winter Meeting, and the 2022 Asia Meeting of the Econometric Society in East and South-East Asia for valuable comments and suggestions.

Data Availability Statement. Replication code and data for this article are available online in Canen and Sugiura (2023) at https://doi.org/10.7910/DVN/VLMMZQ.

Supplementary Material. For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2023.40.

Competing Interest. The authors declare none.

References

Acemoglu, D., C. García-Jimeno, and J. A. Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105 (8): 2364–2409.

Anderson, J. E., and E. van Wincoop. 2003. Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93 (1): 170–192.

Aronow, P. M., C. Samii, and V. A. Assenova. 2015. "Cluster–Robust Variance Estimation for Dyadic Data." *Political Analysis* 23 (4): 564–577.

Barabási, A. L., and R. Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286 (5439): 509–512.

Bliss, H., and B. Russett. 1998. "Democratic Trading Partners: The Liberal Connection, 1962–1989." *Journal of Politics* 60 (4): 1126–1147.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics* 29 (2): 238–249.

Cameron, A. C., and D. L. Miller. 2014. "Robust Inference for Dyadic Data." Working Paper.

Canen, N., J. Schwartz, and K. Song. 2020. "Estimating Local Interactions among Many Agents Who Observe Their Neighbors." *Quantitative Economics* 11 (1): 917–956.

Canen, N., and K. Sugiura. 2023. "Replication Data for: Inference in Linear Dyadic Data Models with Network Spillovers." Harvard Dataverse, V1. https://doi.org/10.7910/DVN/VLMMZQ

Chiang, H. D., and B. Y. Tan. 2023. "Empirical Likelihood and Uniform Convergence Rates for Dyadic Kernel Density Estimation." *Journal of Business & Economic Statistics* 41 (3):906–914.

Cranmer, S. J., and B. A. Desmarais. 2016. "A Critique of Dyadic Design." *International Studies Quarterly* 60 (2): 355–362.

Eaton, J., and S. Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70 (5): 1741–1779.

Erdös, P., and A. Rényi. 1959. "On Random Graphs I." *Publicationes Mathematicae* 6: 290–297.

Erdös, P., and A. Rényi. 1960. "On the Evolution of Random Graphs." *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5: 17–61.

Erikson, R. S., P. M. Pinto, and K. T. Rader 2014. "Dyadic Analysis in International Relations: A Cautionary Tale." *Political Analysis* 22 (4): 457–463.

Gowa, J., and S. Y. Kim. 2005. "An Exclusive Country Club: The Effects of the Gatt on Trade, 1950–94." *World Politics* 57 (4): 453–478.

Graham, B. S. 2020a. "Chapter 2: Dyadic Regression." In *The Econometric Analysis of Network Data*, 1st Edn., edited by B. Graham and Á. d. Paula, 23–40. Cambridge: Academic Press.

Graham, B. S. 2020b. "Chapter 2: Network Data." In *Handbook of Econometrics, Volume 7A*, edited by S. N. Durlauf, L. P. Hansen, J. J. Heckman, and R. L. Matzkin, Volume 7 of Handbook of Econometrics, 111–218. Amsterdam: Elsevier.

Graham, B. S., F. Niu, and J. L. Powell. 2022. "Kernel Density Estimation for Undirected Dyadic Data." Working Paper.

Hansen, B. E., and S. Lee. 2019. "Asymptotic Theory for Clustered Samples." *Journal of Econometrics* 210 (2): 268–290.

Harmon, N., R. Fisman, and E. Kamenica. 2019. "Peer Effects in Legislative Voting." *American Economic Journal: Applied Economics* 11 (4): 156–180.
Helpman, E., M. Melitz, and Y. Rubinstein. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." *Quarterly Journal of Economics* 123 (2): 441–487.
Hoff, P. D., and M. D. Ward. 2004. "Modeling Dependencies in International Relations Networks." *Political Analysis* 12 (2): 160–175.
Kelejian, H. H., and I. R. Prucha. 2010. "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances." *Journal of Econometrics* 157 (1): 53–67.
Kojevnikov, D., V. Marmer, and K. Song. 2021. Limit Theorems for Network Dependent Random Variables. *Journal of Econometrics* 222 (2): 882–908.
Leung, M. P., and H. R. Moon. 2021. "Normal Approximation in Large Network Models." Working Paper.
Lowe, M., and D. Jo. 2021. "Legislature Integration and Bipartisanship: A Natural Experiment in Iceland." CESifo Working Paper.
Lustig, H., and R. J. Richmond. 2020. "Gravity in the Exchange Rate Factor Structure." *Review of Financial Studies* 33 (8): 3492–3540.
Mansfield, E. D., H. V. Milner, and B. P. Rosendorff. 2000. "Free to Trade: Democracies, Autocracies, and International Trade." *American Political Science Review* 94 (2): 305–321.
Melitz, M. J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.
Minhas, S., P. D. Hoff, and M. D. Ward. 2019. "Inferential Approaches for Network Analysis: Amen for Latent Factor Models." *Political Analysis* 27 (2): 208–222.
Minhas, S., et al. 2022. "Taking Dyads Seriously." *Political Science Research and Methods* 10 (4): 703–721.
Poast, P. 2016. "Dyads Are Dead, Long Live Dyads! The Limits of Dyadic Designs in International Relations Research." *International Studies Quarterly* 60 (2): 369–374.
Saia, A. 2018. "Random Interactions in the Chamber: Legislators' Behavior and Political Distance." *Journal of Public Economics* 164: 225–240.
Tabord-Meehan, M. 2019. "Inference with Dyadic Data: Asymptotic Behavior of the Dyadic-Robust t-Statistic." *Journal of Business & Economic Statistics* 37 (4): 671–680.

Cite this article: Canen, N. and Sugiura, K. (2023). Inference in Linear Dyadic Data Models with Network Spillovers. *Political Analysis*, 1–18. https://doi.org/10.1017/pan.2023.40