Robust Inference in Locally Misspecified Bipartite Networks*

Luis E. Candelaria[†] Yichong Zhang[‡]

November 2023

Link to latest version

Abstract

This paper introduces a methodology to conduct robust inference in bipartite networks under local misspecification. We focus on a class of dyadic network models with misspecified conditional moment restrictions. The framework of misspecification is local as the magnitude of misspecification vanishes as the sample size grows. We utilize this local asymptotic approach to construct a minimax optimal robust estimator that attains the shortest confidence interval length within a neighborhood of misspecification. Additionally, we introduce bias-aware confidence intervals that account for the effect of the local misspecification. These confidence intervals are asymptotically valid for the structural parameter of interest under sparse network asymptotics, both in the correctly specified and locally misspecified case. Monte Carlo experiments demonstrate that the robust estimator performs well in finite samples and sparse networks. As an empirical illustration, we study the formation of a scientific collaboration network among economists.

*We are thankful to Stéphane Bonhomme, Áureo de Paula, Michal Kolesár for helpful comments and suggestions. Yichong Zhang acknowledges the financial support from the NSFC under the grant No. 72133002. All remaining errors are our own.

[†]University of Warwick. E-mail: l.candelaria@warwick.ac.uk.

[‡]Singapore Management University. E-mail: yczhang@smu.edu.sg.

1 Introduction

Bipartite networks are embedded in a wide range of economic interactions that entail bilateral relations, for example, those between exporters and importers, buyers and sellers, and scientists and research teams. Dyadic regression with two-sided heterogeneity represents a leading approach for analyzing bipartite networks as it can adequately account for dyad dependence and link sparsity (Graham 2020). While it yields valid econometric inference for a structural parameter of interest under correct model specification, its performance under local misspecification has not been explored. To the best of our knowledge, this is the first paper to study the effects of local misspecification on a dyadic regression model for bipartite networks.

In this paper, we develop a methodology to conduct robust inference in bipartite networks under local misspecification. We focus on a class of dyadic models characterized by conditional moment restrictions when the model misspecification prevents these restrictions from holding exactly. Some situations that lead to the failure of these moment restrictions in dyadic network models include incorrect specification of parametric assumptions, latent homophily, and strategic linking decisions. Our framework is one of local misspecification as the magnitude of misspecification vanishes as the sample size grows. This modelling approach is intended as an asymptotic device, which is particularly useful to provide a tractable approximation for the effect of the model misspecification. Moreover, it enables us to perform inference on a structural parameter of interest rather than on a pseudo-true parameter.

We propose a robust estimator that accounts for local misspecification on the conditional moment restrictions. To construct the robust estimator, we consider a sieve approximation to form unconditional moment restrictions that grow in number with the sample size. The robust estimator is then computed as a one-step estimator, which improves on an initial estimator that ignores the model misspecification by incorporating an adjustment term that minimizes the sensitivity of the inferential procedure to local deviations in the unconditional moment restrictions. In particular, this adjustment term is chosen optimally to minimize the confidence interval length. As a consequence, the robust estimator is minimax optimal in the sense that given the worst-case bias in a neighborhood prespecified by the researcher, the robust estimator attains the shortest confidence interval length within a class of regular estimators (Newey 1990).

The resulting confidence intervals are 'honest', as they are widened to account for the worst-case bias induced by the local misspecification within a prespecified neighborhood (Donoho 1994; Armstrong and Kolesár 2021; Bonhomme and Weidner 2022). This construction ensures that the bias-aware confidence intervals contain the structural parameter of interest under both correct specification and local misspecification with a prespecified coverage probability. The neighborhood of misspecification is determined ex-ante by the researcher but can be adjusted to conduct sensitivity analysis on the structural parameter of interest. Additionally, we show that in our setting, the misspecification neighborhood takes on a simple form, which enables us to obtain a closed-form expression for the adjustment term. We describe in detail the construction of the robust estimator and bias-aware confidence intervals through two examples of empirical interest: (i) when the parameter of interest corresponds to a homophily parameter that indexes the dyadic regression, and (ii) when the parameter of interest is equal to the average out-degree of the network.

To illustrate the performance of our methodology in an empirical application, we study the formation of a scientific collaboration network among economists. While collaboration among scientists is pivotal for research productivity and innovation, the factors that determine the creation of these connections have remained underexplored (Goyal, Van Der Leij, and Moraga-González 2006; Anderson and Richards-Shubik 2022; Hsieh, Konig, Liu, and Zimmermann 2022). In addition, the existing approaches used to analyze these collaborations rely on strong parametric assumptions or limit the degree of unobserved heterogeneity, making them susceptible to model misspecification. To address this gap, we implement our robust methodology to study the factors driving the formation of scientific collaborations among economists. The web of collaborations is described as a bipartite network connecting scientists and research projects, and it is constructed using the universe of published papers in the top American economic journals of general interest during the period of 2000 to 2006. We describe in detail the dataset and network construction in Section 7.

As parameters of interest, we focus on those capturing the assortative matching among scientists and the attributes of a research project. Additionally, we estimate global features of the network, such as the average out-degree. Our analysis documents that homophily in the research field of expertise is a strong predictor for link formation. Moreover, on average, highly productive scientists are more likely to collaborate on higher-impact projects, measured by aggregate citations. The results also show that, during the sample period, female authors are more likely to sort in teams with a higher share of female authors. Finally, the evidence shows that our methodology leads to a significant reduction in the confidence intervals' length.

In the rest of this section, we relate the contributions of this paper to the existing literature. Our paper is related to the literature on robust statistics (Huber 1964; Huber and Ronchetti 2009; Rieder 1994), sensitivity analysis (Leamer 1985; Imbens 2003; Chen, Tamer, and Torgovitsky 2011; Nevo and Rosen 2012; Masten and Poirier 2020, 2021), and local misspecification (Newey 1985; Conley, Hansen, and Rossi 2012; Bugni, Canay, and Guggenberger 2012; Kitamura, Otsu, and Evdokimov 2013; Andrews, Gentzkow, and Shapiro 2017, 2020; Bugni and Ura 2019; Armstrong and Kolesár 2021; Bonhomme and Weidner 2022; Armstrong, Weidner, and Zeleneev 2022; Armstrong, Kline, and Sun 2023; Christensen and Connault 2023). The most closely related papers to ours are those by Andrews et al. (2017); Armstrong and Kolesár (2021); Bonhomme and Weidner (2022) and Christensen and Connault (2023), who rely on local misspecification to conduct sensitivity analysis. In contrast to these papers, this work focuses on a class of models characterized by conditional moment restrictions, which is not nested within their settings. Moreover, the object of study is entirely different as we analyze dyadic models for bipartite networks. On a technical side, bipartite network exhibit patterns of dyad dependence that preclude us from using efficiency results invoked by the previous papers (cf. Kitamura et al. 2013; Bonhomme and Weidner 2022).

The paper also contributes to the literature on dyadic regression (Tabord-Meehan 2019; Davezies, D'Haultfœuille, and Guyonvarch 2021; Menzel 2021; Graham 2022). The closest paper to ours is the one by Graham (2022). In this insightful paper, Graham derives the asymptotic properties of a logistic regression under sparse network asymptotics and shows that under global misspecification, the pseudo-true parameter solves a particular Kullback–Leibler Information Criterion minimization. In contrast, we construct a one-step estimator for a dyadic regression under local misspecification, which allows for more general forms of misspecification. Moreover, our methodology also differs substantially as we make use of the local approximation to conduct inference on the true structural parameter rather than a pseudo-true parameter.

The rest of the paper is organized as follows. In Section 2, we introduce dyadic model for bipartite networks under local misspecification. In Section 3, we introduce the robust estimator and the bias-aware confidence intervals. We present the main theoretical results in Section 4. Section 5 illustrates our methodology in two examples of empirical interest. Section 6 reports evidence from Monte Carlo experiments. In Section 7, we present an empirical application, and Section 8 concludes. Additional results and proofs are collected in the appendices.

2 Model

We consider a bipartite network of scientific collaborations with N scientists and M projects. For any $i \in [N]$, let X_i and A_i denote vectors of scientist i-specific observed and unobserved attributes. Similarly, for any $j \in [M]$, let W_j and B_j denote vectors of project j-specific observed and unobserved attributes. For any pair $i \in [N]$ and $j \in [M]$, U_{ij} denotes a scientist-project-specific unobserved attribute. Let $\{X_i, A_i\}_{i=1}^{N}$, $\{W_j, B_j\}_{j=1}^{M}$, $\{U_{ij}\}_{i\in[N],j\in[M]}$ be sequences of i.i.d. random variables that are independent of each other. Also, define $n \equiv N+M$, where N and M are assumed to grow at the same rate as in Assumption 1.

For any finite n, the X-W-exchangeable graphon is defined as

$$Y_{ij} = h_n\left(v_{ij}, X_i, W_j, A_i, B_j, U_{ij}\right), \tag{2.1}$$

where v_{ij} represents an i.i.d. latent mixing factor, as discussed in Graham (2022), and h_n is a measurable function that maps $(v_{ij}, X_i, W_j, A_i, B_j, U_{ij}) \mapsto \{0, 1\}$. Equation (2.1) is used as a nonparametric data generating process for the bipartite network of scientific collaborations with N scientists and M projects. Intuitively, equation (2.1) states that $Y_{ij} = 1$ if scientist i collaborates in project j, and 0 otherwise. Moreover, this decision depends on both scientist i's and project j's observed and unobserved attributes.

We assume that the conditional probability of establishing a collaboration link is given by

$$\Pr\left[Y_{ij} = 1 \mid X_i, W_j\right] = \int e\left(\alpha_{0,n} + Z_{ij}^{\top}\beta_0 + n^{-1/4}v\right)\pi(v)dv, \tag{2.2}$$

where $e(u) = \exp(u)/(1+\exp(u))$ is the standard logistic CDF, $Z_{ij} = z(X_i, W_j) \in \Re^{d_z}$ is a vector-valued distance of the observed attributes X_i and W_j, and $\pi(\cdot)$ is a proper PDF with mean zero, $\int v\pi(v)dv = 0$, and finite variance, $\int v^2\pi(v)dv = \sigma^2$.

The coefficient $\alpha_{0,n}$ is allowed to vary with the sample size to accommodate for sparse network asymptotics. In particular, we consider $\alpha_{0,n} \equiv \log(\alpha_0/\log(n))$, which jointly with the specification of the conditional distribution in equation (2.2) and Assumption 1, ensure that in the limit the network is sparse. Consequently, the average in-degree and out-degree of this network will be bounded. That is, for any $i \in [N]$ and $j \in [M]$, let $\rho_n \equiv \mathbb{E}_n\left[Y_{ij}\right]$ denote the marginal probability of forming a link, then $\lim_{n\to\infty} M\rho_n = \lambda_0^a < \infty$ and $\lim_{n\to\infty} N\rho_n = \lambda_0^p < \infty$.

The linear index $Z_{ij}^{\top}\beta_0$ in equation (2.2) captures the contribution that sharing similar observed attributes has on a scientist i's decision to collaborate in the project j. In other words, this component represents an assortative matching mechanism for the scientific collaboration network. Consequently, the coefficient β_0 is interpreted as a homophily parameter.

Finally, notice that the latent component v affects the conditional probability of establishing scientific collaboration. In this setting, the component $n^{-1/4}v$ represents a source of local misspecification.[1] The misspecification design is local as it is indexed by n and vanishes away as the sample size increases at a rate of $n^{-1/4}$. In Section 3, we show that at this rate, the local misspecification induces a bias term into the asymptotic distribution of the parameter of interest. Moreover, the magnitude of the misspecification, and thus, of the asymptotic bias, is determined by the variance σ^2 of the distribution of the latent factor v.

We now provide three relevant empirical illustrations that can be analyzed using the local misspecification framework characterized by equation (2.2).

Example 1 (Latent homophily)**.** *Consider the case in which scientist i's decision to participate in project j is given by*

$$Y_{ij} = F\left[\alpha_{0,n} + Z_{ij}^{\top}\beta_0 + A_i + B_j + n^{-1/4}v_{ij} + U_{ij}\right]$$

where F represents a known function increasing function (e.g., $F(U) = sign(u)$), $v_{ij} = \psi(v_i, v_j)$

[1]The coefficient associated with v is normalized to one without loss of generality. However, notice that it can be set to be any known function of the observed attributes.

with v_i and v_i being i.i.d. unobserved fixed effects and $\psi(\cdot,\cdot)$ some finite known function (e.g., $v_{ij} = |v_i - v_j|^\delta$ for $\delta \geq 1$), and $A_i + B_j + U_{ij}$ has CDF given by $e(\cdot)$. In this setting, A_i and B_j represent additive fixed effects, while v_{ij} captures unobserved homophily on the latent components v_i and v_j.

Example 2 (Semiparametric distribution)**.** *Alternatively, consider the nonlinear dyadic regression*

$$Y_{ij} = \mathbf{1}\left[\alpha_{0,n} + Z_{ij}^\top \beta_0 \geq n^{-1/4} v_{ij} + U_{ij}\right],$$

where U_{ij} has a logistic CDF and v_{ij} represents a mixing latent component with an unknown distribution that has mean zero with finite variance. In this setting, the composite error term $n^{-1/4}v_{ij} + U_{ij}$ will have an unknown distribution rendering the estimation of $(\alpha_{0,n}, \beta_0)$ nonparametric.

Example 3 (Network Externalities)**.** *Suppose that the graphon $[Y_{ij}]_{i\in[N],j\in[M]}$ is formed according to the strategic network formation model with incomplete information as described in Leung (2015); Ridder and Sheng (2020). Then, under regularity conditions, the Bayesian Nash Equilibrium Strategy is given by*

$$Y_{ij} = \mathbf{1}\left[\alpha_{0,n} + Z_{ij}^\top \beta_0 + n^{-1/4} v_{ij} \left(\mathbb{E}\left[\mathcal{E}_{ij} \mid X_N, W_M, \mu\right]\right) \geq U_{ij}\right]$$

with $\mathcal{E}_{ij} \equiv \mathcal{E}_{ij}[Y_n, X_N, W_M]$, $Y_n = [Y_{ij}]_{i\in[N],j\in[M]}$, $X_N = (X_1, \cdots, X_N)$, $W_M = (W_1, \cdots, W_M)$, and μ denotes the equilibrium belief profile. The component $\mathcal{E}_{ij}$ represents an endogenous network externality. Some examples include: (i) In-degree of project j

$$\frac{1}{N}\sum_{k\neq i} Y_{kj}\omega(X_i, X_k)$$

where $\omega(X_i, X_k)$ represents a weighting function. (ii) Out-degree of scientist i

$$\frac{1}{M}\sum_{k\neq j} Y_{ik}\omega(W_j, W_k).$$

where $\omega(W_j, W_k)$ represents a weighting function. (iii) Transitivity effect

$$\mathbf{1}\left[\sum_{k\neq i}\sum_{l\neq j} Y_{kl}Y_{il} \geq 0\right],$$

which captures the utility from collaborating in project j with a coauthor k who has collaborated in at least other project $l \neq k$.

As parameter of interest, we focus on the scalar $\Psi_{0,n}$ that is defined as

$$\Psi_{0,n} \equiv \int \mathbb{E}\gamma_n \left(D_{ij}, \theta_{0,n}, n^{-1/4}v\right) \pi(v)dv, \tag{2.3}$$

where γ_n is a measurable function that maps $\left(D_{ij}, \theta_{0,n}, n^{-1/4}v\right) \mapsto \Re$, with $D_{ij} \equiv (Y_{ij}, Z_{ij}^\top)^\top$ and $\theta_{0,n} \equiv (\alpha_{0,n}, \beta_0)$. Notice that the parameter of interest $\Psi_{0,n}$ is allowed to be misspecified itself (cf. Armstrong and Kolesár 2021). Next, we provide some examples of the parameter of interest as described by equation (2.3).

Example 4 (Individual Parameters)**.** *Suppose that we are interested in doing inference on an individual element of the homophily parameter β_0 or the intercept parameter $\alpha_{0,n}$, then the parameter of interest can be defined as $\Psi_{0,n} = \theta_{0,n}^{(k)}$, where $\theta_{0,n}^{(k)}$ denotes the kth element in $\theta_{0,n} \in \Re^{d_z+1}$.*

Example 5 (Average in and out-Degree)**.** *We might be interested in doing inference on the average number of scientists that participate in a randomly selected project. Then, the parameter $\Psi_{0,n}$ can be defined as the average in-degree of the network, i.e,*

$$\Psi_{0,n} = N \int \mathbb{E}\left[e\left(\alpha_{0,n} + Z_{ij}^\top \beta_0 + n^{-1/4}v\right)\right] \pi(v)dv.$$

Alternatively, we might be interested in doing inference on the average number of collaborations in which a randomly selected scientist participates, then $\Psi_{0,n}$ can be defined as the average out-degree of the network, i.e,

$$\Psi_{0,n} = M \int \mathbb{E}\left[e\left(\alpha_{0,n} + Z_{ij}^\top \beta_0 + n^{-1/4}v\right)\right] \pi(v)dv.$$

Example 6 (Aggregated Marginal Effect)**.** *Suppose that $Z_{ij}^{(1)}$ is continuous. The parameter of interest $\Psi_{0,n}$ could be defined as the aggregated marginal effect of changing $Z_{ij}^{(1)}$, i.e.,*

$$\Psi_{0,n} \quad = \quad n \int \mathbb{E}\frac{\partial}{\partial Z_{ij}^{(1)}} e\left(\alpha_{0,n} + Z_{ij}^\top \beta_0 + n^{-1/4}v\right) \pi(v)dv.$$

Example 7 (Average Partial Effect)**.** *Alternatively, we might be interested in doing inference on the average partial effect, then $\Psi_{0,n}$ can be defined as*

$$\Psi_{0,n} = n \int \mathbb{E}\frac{\partial}{\partial \beta} e\left(\alpha_{0,n} + Z_{ij}^\top \beta_0 + n^{-1/4}v\right)\Big|_{\beta=\beta_0} \pi(v)dv.$$

2.1 Robust Confidence Interval

The robust estimator $\hat{\Psi}(\hat{\kappa})$ of $\Psi_{0,n}$ is defined as the following one-step regular estimator

$$\hat{\Psi}(\hat{\kappa}) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\gamma_n(D_{ij}, 0, \hat{\theta}_{\text{initial}}) + \hat{\kappa}^\top \left[\frac{n}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(Y_{ij} - e(R_{ij}^\top \hat{\theta}_{\text{initial}}))H_{ij}\right], \tag{2.4}$$

where $\hat{\theta}_{\text{initial}}$ is a $\sqrt{n}$-consistent initial estimator of $\theta_{0,n}$, i.e., $\sqrt{n}\left(\hat{\theta}_{\text{initial}} - \theta_{0,n}\right) = O_p(1)$, H_{ij} is a k_n-dimensional vector consists of sieve basis functions of X_i and W_j, and $\hat{\kappa}$ is a k_n-dimensional estimator of the sensitivity parameter κ. In practice, we use the logistic MLE as our initial estimator, i.e.,

$$\hat{\theta}_{\text{initial}} = (\hat{\alpha}_n, \hat{\beta}_n) \equiv \arg\max_{\theta} L_n(\theta),$$

where

$$L_n(\theta) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} l_{ij}(\theta), \quad l_{ij}(\theta) = Y_{ij}R_{ij}^\top\theta - \log(1 + \exp(R_{ij}^\top\theta)).$$

The first component in equation (2.4) is a plug-in estimator of $\Psi_{0,n}$ evaluated at the initial estimator. Under local misspecification, the asymptotic distribution of the plug-in estimator will be will not be centered at $\Psi_{0,n}$ due to the asymptotic bias induced by $n^{-1/4}v$.

The second component in equation (2.4) corresponds to a one-step adjustment that accounts for the asymptotic bias generated by the local misspecification. The term in brackets represents the sample analogue of the moment restrictions in

$$\mathbb{E}\left[Y_{ij} - e\left(R_{ij}^\top\theta\right) H_{ij}\right] = 0, \tag{2.5}$$

which holds with equality to zero when evaluated at $\theta_{0,n}$ in the absence of local misspecification. The sample analogue in (2.4) is scaled by n to account for the sparse network asymptotics (cf. Graham 2022). The vector $\hat{\kappa}$ is an estimate of κ, which captures the response of the estimator $\hat{\Psi}(\hat{\kappa})$ to the misspecification of the $k_n \times 1$ moment conditions in (2.5). Thus, following Andrews et al. (2020); Armstrong and Kolesár (2021); Bonhomme and Weidner (2022), we refer to κ as a sensitivity parameter. In Section 2.3, we show how to optimally choose κ for this setting.

In contrast to the existing literature on sensitivity analysis that relies on maximum likelihood methods or method of moments, the dimension of the sensitivity parameter κ grows with the dimension of the sieve space. This is an implication that the underlying moment restrictions in equation (2.5) are constructed based on the conditional moment restriction

$$\mathbb{E}\left[Y_{ij} - e\left(R_{ij}^{\top}\theta\right) \mid X_i, W_j\right] = 0,$$

which holds with equality when evaluated at $\theta_{0,n}$ in the absence of local misspecification. Notice, then that the growing sieve space will determine the limiting distribution of the one-step robust estimator.

Next, we describe the computation of the bias-aware confidence inference that will be central to performing robust inference on $\Psi_{0,n}$. First, following Newey (1990) and Armstrong and Kolesár (2021), we require our one-step robust estimator to be regular, which given the sensitivity parameter κ, it can be represented as

$$G^{\top}\kappa = \Gamma, \tag{2.6}$$

where Γ denotes the derivative of the $\Psi_{0,n}$ with respect to θ and G denotes the score of the moment restriction in equation (2.5) with respect to $\theta_{0,n}$. Formal definitions of Γ and G are provided in Assumption 2. Under this requirement, provided that $\hat{\kappa}$ is a consistent estimator of κ, i.e., $||\hat{\kappa} - \kappa||_2 = o_P(1)$ and the initial estimator is $\sqrt{n}$-consistent, i.e., $\hat{\theta}_{\text{initial}} - \theta_{0,n} = O_P(n^{-1/2})$, we have a strong approximation of the distribution of the robust estimator $\hat{\Psi}(\hat{\kappa})$ which does not depend on asymptotic distribution of the initial estimator:

$$\sqrt{n}(\hat{\Psi}(\hat{\kappa}) - \Psi_{0,n}) - \omega_n(\kappa) = o_P(1), \tag{2.7}$$

where

$$\omega_n(\kappa) \overset{d}{=} \mathcal{N}\left(\sigma^2\left[\Delta_v + \kappa^{\top}B_{\theta}\right], \Sigma_{\gamma} + 2\Sigma_{H,\gamma}^{\top}\kappa + \kappa^{\top}\Sigma_H\kappa\right).$$

The asymptotic bias is comprised of two leading components. The first element $\sigma^2\Delta_v$, with $\Delta_v \equiv \frac{1}{2}\lim_{n\to\infty}\mathbb{E}\partial_{vv}\gamma_n(D_{ij}, v, \theta_{0,n})|_{v=0}$, represents the effect that the local misspecification has on the initial plug-in estimator of the robust estimator. In other words, it is the asymptotic bias induced by the first term in equation (2.4). The second element $\sigma^2\left[\kappa^{\top}B_{\theta}\right]$ with $B_{\theta} = \frac{1}{2}\lim_{n\to\infty}\mathbb{E}\left[e\left(R_{ij}^{\top}\theta_{0,n}\right)\left(e\left(R_{ij}^{\top}\theta_{0,n}\right) - 1\right)\left(1 - 2e\left(R_{ij}^{\top}\theta_{0,n}\right)\right)H_{ij}\right]$, represents the asymptotic bias due to the local misspecification of using $e(R_{ij}^{\top}\theta)$ as $\mathbb{P}(Y_{ij} = 1|X_i, W_j)$ in the unconditional moment restrictions in equation (2.5). In other words, it is the asymptotic bias induced by the second term in equation (2.4).

Regarding the asymptotic variance, $(\Sigma_H, \Sigma_{H,\gamma}, \Sigma_{\gamma})$ are three submatrices of a $(k_n+1)\times(k_n+1)$ covariance matrix $\Sigma = \begin{pmatrix}\Sigma_H & \Sigma_{H,\gamma}\\ \Sigma_{H,\gamma}^{\top} & \Sigma_{\gamma}\end{pmatrix}$, and Σ is formally defined in the next section. Here, we referred to $\omega_n(\kappa)$ as the strong approximation rather than the limit under weak convergence because, although deterministic, the dimensions of κ, Σ_H, $\Sigma_{H,\gamma}$, and thus themselves, implicitly depend on the sample size n.

Notice that the asymptotic bias of the one-step robust estimator $\hat{\Psi}(\hat{\kappa})$ depends on the magnitude σ^2 of the local misspecification. Because σ^2 is unidentified, it renders the asymptotic distribution in equation (2.7) not immediately useful to conduct inference on $\Psi_{0,n}$. Nevertheless, if $\overline{\sigma}^2 \in \Re_+$ is an upper bound for σ^2, we can define the worst-case bias within the misspecification neighborhood $\mathcal{M} \equiv \left\{\sigma^2 \in \Re_+ : \sigma^2 \leq \overline{\sigma}^2\right\}$ as

$$\mathcal{B}_M(\kappa) \equiv \sup_{\sigma^2 \in \mathcal{M}} \left|\sigma^2 \left[\Delta_v + \hat{\kappa}^\top B_\theta\right]\right| = \overline{\sigma}^2 \left|\Delta_v + \kappa^\top B_\theta\right|.$$

Then, the infeasible bias-aware confidence interval (CI) for $\Psi_{0,n}$ is defined as

$$CI_{1-\alpha}(\kappa) \equiv \hat{\Psi}(\kappa) \pm CV_\alpha \left(\frac{\mathcal{B}_M(\kappa)}{(\Sigma_\gamma + 2\Sigma_{H,\gamma}^\top \kappa + \kappa^\top \Sigma_H \kappa)^{1/2}}\right) \frac{(\Sigma_\gamma + 2\Sigma_{H,\gamma}^\top \kappa + \kappa^\top \Sigma_H \kappa)^{1/2}}{\sqrt{n}}, \tag{2.8}$$

where $CV_\alpha(t)$ denotes the $1-\alpha$ quantile of $|\mathcal{Z}|$ where $\mathcal{Z} \sim \mathcal{N}(t,1)$. Given consistent estimators $\hat{\kappa}$, $\hat{\Delta}_v$, $\hat{B}_\theta$, $\hat{\Sigma}_H$, $\hat{\Sigma}_{H,\gamma}$, and $\hat{\Sigma}_2$ of their population counterparts, the feasible bias-aware confidence interval for $\Psi_{0,n}$ is defined as

$$\widehat{CI}_{1-\alpha}(\hat{\kappa}) \equiv \hat{\Psi}(\hat{\kappa}) \pm CV_\alpha \left(\frac{\hat{\mathcal{B}}_M(\hat{\kappa})}{(\hat{\Sigma}_\gamma + 2\hat{\Sigma}_{H,\gamma}^\top \hat{\kappa} + \hat{\kappa}^\top \hat{\Sigma}_H \hat{\kappa})^{1/2}}\right) \frac{(\hat{\Sigma}_\gamma + 2\hat{\Sigma}_{H,\gamma}^\top \hat{\kappa} + \hat{\kappa}^\top \Sigma_H \hat{\kappa})^{1/2}}{\sqrt{n}}, \tag{2.9}$$

where

$$\hat{\mathcal{B}}_M(\hat{\kappa}) = \overline{\sigma}^2 \left[\hat{\Delta}_v + \hat{\kappa}^\top \hat{B}_\theta\right].$$

2.2 Variance and Bias Components

In this section, we propose the estimators $\hat{\Delta}_v$, $\hat{B}_\theta$, $\hat{\Sigma}_H$, $\hat{\Sigma}_{H,\gamma}$, and $\hat{\Sigma}_2$ of their population counterparts. Recall that

$$\Sigma = \begin{pmatrix} \Sigma_H & \Sigma_{H,\gamma} \\ \Sigma_{H,\gamma}^\top & \Sigma_\gamma \end{pmatrix}.$$

The matrix Σ is defined as the covariance of the joint scores

$$s_{ij} = \begin{pmatrix} n\left(Y_{ij} - \int e(R_{ij}^\top \theta_{0,n} + n^{-1/4} v)\pi(v)dv\right) H_{ij} \\ \gamma_n(D_{ij}, 0, \theta_{0,n}) - \mathbb{E}\gamma_n(D_{ij}, 0, \theta_{0,n}) \end{pmatrix} \equiv (s_{Y,ij} H_{ij}^\top, s_{\gamma,ij})^\top.$$

Specifically, let $\overline{s}_{ij} = \mathbb{E}(s_{ij}|X_i, A_i, W_j, B_j)$, $\overline{s}^a_{1i} = \mathbb{E}(\overline{s}_{ij}|X_i, A_i)$, $\overline{s}^p_{1j} = \mathbb{E}(\overline{s}_{ij}|W_j, B_j)$, $\Sigma^a_n = \mathbb{E}\overline{s}^a_{1i}(\overline{s}^a_{1i})^\top$, $\Sigma^p_n = \mathbb{E}\overline{s}^p_{1j}(\overline{s}^p_{1j})^\top$, and $\Sigma_{2n} = \mathbb{E}(s_{ij} - \overline{s}_{ij})(s_{ij} - \overline{s}_{ij})^\top$. Because $D_{ij} = (X_i^\top, W_j^\top)$, we have that

$$\overline{s}_{\gamma,ij} \equiv \mathbb{E}(s_{\gamma,ij}|X_i, A_i, W_j, B_j) = s_{\gamma,ij}.$$

Then, the covariance matrix Σ is given by

$$\Sigma \equiv \begin{pmatrix} \Sigma_H & \Sigma_{H,\gamma} \\ \Sigma^\top_{H,\gamma} & \Sigma_\gamma \end{pmatrix} = \frac{\Sigma^a_n}{1-\phi} + \frac{\Sigma^p_n}{\phi} + \frac{\Sigma_{2n}}{\phi(1-\phi)}$$

with

$$\Sigma^a_n = \mathbb{E}\begin{pmatrix} [\mathbb{E}(s_{Y,ij}H_{ij}|X_i, A_i)][\mathbb{E}(s_{Y,ij}H_{ij}|X_i, A_i)]^\top & \mathbb{E}(s_{Y,ij}H_{ij}|X_i, A_i)\mathbb{E}(s_{\gamma,ij}|X_i, A_i) \\ \mathbb{E}(s_{Y,ij}H^\top_{ij}|X_i, A_i)\mathbb{E}(s_{\gamma,ij}|X_i, A_i) & [\mathbb{E}(s_{\gamma,ij}|X_i, A_i)]^2 \end{pmatrix},$$

$$\Sigma^p_n = \mathbb{E}\begin{pmatrix} [\mathbb{E}(s_{Y,ij}H_{ij}|W_j, B_j)][\mathbb{E}(s_{Y,ij}H_{ij}|W_j, B_j)]^\top & \mathbb{E}(s_{Y,ij}H_{ij}|W_j, B_j)\mathbb{E}(s_{\gamma,ij}|W_j, B_j) \\ \mathbb{E}(s_{Y,ij}H^\top_{ij}|W_j, B_j)\mathbb{E}(s_{\gamma,ij}|W_j, B_j) & [\mathbb{E}(s_{\gamma,ij}|W_j, B_j)]^2 \end{pmatrix},$$

and

$$\Sigma_{2n} = \mathbb{E}\begin{pmatrix} Var(s_{Y,ij}|X_i, W_j)H_{ij}H^\top_{ij} & 0 \\ 0 & 0 \end{pmatrix}.$$

2.3 Sensitivity Parameter

In this section, we describe the definition and estimation of the optimal sensitivity parameter κ^* and its estimation. Recall the infeasible bias-aware CI for $\Psi_{0,n}$ with sensitivity parameter κ is $CI_{1-\alpha}(\kappa)$ defined in equation (4.1). The asymptotic length for $CI_{1-\alpha}(\kappa)$ is denoted as $L(\kappa)$ and defined as

$$L(\kappa) = 2CV_\alpha\left(\frac{\mathcal{B}_M(\kappa)}{(\Sigma_\gamma + 2\Sigma^\top_{H,\gamma}\kappa + \kappa^\top\Sigma_H\kappa)^{1/2}}\right)(\Sigma_\gamma + 2\Sigma^\top_{H,\gamma}\kappa + \kappa^\top\Sigma_H\kappa)^{1/2}.$$

We optimally choose κ as the parameter value that attains the shortest confidence interval length, i.e.,

$$\kappa^* = \arg\min_\kappa L(\kappa)$$

subject to the regularity condition given by equation (2.6). Following Armstrong and Kolesár (2021), we can convert this optimization problem into a minimization of the asymptotic variance subject to a constraint on the worst-case bias and the regularity condition in (2.6). That is,

$$\min_{\kappa} \Sigma_\gamma + 2\Sigma_{H,\gamma}\kappa + \kappa^\top \Sigma_H \kappa \quad \text{s.t.} \quad \sup_{\sigma^2 \in \mathcal{M}} \left| \sigma^2 \left[\Delta_v + \kappa^\top B_\theta \right] \right| \leq \overline{B} \quad \text{and} \quad G^\top \kappa = \Gamma. \tag{2.10}$$

Let λ denote the Lagrange multiplier associated with the worst-case bias constraint; then we can cast the optimization problem in (2.10) as the following equivalent convex optimization problem

$$\begin{aligned} &\min_{\kappa} \left(\Sigma_\gamma + \lambda \overline{\sigma}^2 \Delta_v^2 \right) + 2 \left[\Sigma_{H,\gamma} + \lambda M \Delta_v B_\theta^\top \right] \kappa + \kappa^\top \left\{ \Sigma_H + \lambda \overline{\sigma} B_\theta B_\theta^\top \right\} \kappa \\ &\text{s.t.} \quad G^\top \kappa = \Gamma, \end{aligned} \tag{2.11}$$

where $\lambda \in [0, \infty)$ is the Lagrangian multiplier for the constraint on the worst-case bias. That is, for the constraint $\overline{\sigma}^2 |\Delta_v + \kappa^\top B_\theta| \leq \overline{B}$. By solving (2.11), we obtain the optimal sensitivity parameter $k^*(\lambda)$ as a function of λ defined as

$$\kappa^*(\lambda) \quad = \quad -\Upsilon_\lambda^{-1} \left\{ A_\lambda^\top - G \left(G^\top \Upsilon_\lambda^{-1} G^\top \right)^{-1} \left[\Gamma + G^\top \Upsilon_\lambda^{-1} A_\lambda^\top \right] \right\}, \tag{2.12}$$

where $\Upsilon_\lambda = \left[\Sigma_H + \lambda \overline{\sigma}^2 B_\theta B_\theta^\top \right]$ and $A_\lambda = \left[\Sigma_{H,\gamma} + \lambda \overline{\sigma}^2 \Delta_v B_\theta^\top \right]$. Then, the final κ^* is defined as $\kappa^*(\lambda^*)$, where λ^* is defined as

$$\lambda^* = \arg\min_{\lambda > 0} L(\kappa^*(\lambda)). \tag{2.13}$$

Our estimator of κ^* is $\hat{\kappa}^*(\hat{\lambda}^*)$, where $\hat{\kappa}(\lambda)$ is the sample analog of $\kappa^*(\lambda)$ defined in (2.12) and $\hat{\lambda}^*$ is a consistent estimator of λ^*. Specifically, we define

$$\hat{\kappa}^*(\lambda) = -\widehat{\Upsilon}_\lambda^{-1} \left\{ \widehat{A}_\lambda^\top - \widehat{G} \left(\widehat{G}^\top \widehat{\Upsilon}_\lambda^{-1} \widehat{G}^\top \right)^{-1} \left[\widehat{\Gamma} + \widehat{G}^\top \widehat{\Upsilon}_\lambda^{-1} \widehat{A}_\lambda^\top \right] \right\},$$

where $\widehat{\Upsilon}_\lambda = \left[\widehat{\Sigma}_H + \lambda \overline{\sigma}^2 \widehat{B}_\theta \widehat{B}_\theta^\top \right]$ and $\widehat{A}_\lambda = \left[\widehat{\Sigma}_{H,\gamma} + \lambda \overline{\sigma}^2 \widehat{\Delta}_v \widehat{B}_\theta^\top \right]$. As for $\hat{\lambda}^*$, we note that the asymptotic length of $CI_{1-\alpha}(\kappa)$ can be estimated by $\hat{L}(\kappa)$, which is defined as

$$\hat{L}(\kappa) = 2CV_\alpha \left(\frac{\hat{\mathcal{B}}_M(\kappa)}{(\hat{\Sigma}_\gamma + 2\hat{\Sigma}_{H,\gamma}^\top \hat{\kappa} + \kappa^\top \hat{\Sigma}_H \hat{\kappa})^{1/2}} \right) (\hat{\Sigma}_\gamma + 2\hat{\Sigma}_{H,\gamma}^\top \kappa + \kappa^\top \Sigma_H \kappa)^{1/2}.$$

Then, we can define $\hat{\lambda}^*$ as

$$\hat{\lambda}^* = \arg\min_{\lambda > 0} \hat{L}(\hat{\kappa}^*(\lambda)).$$

Consequently, for a fixed $\alpha \in (0, 1)$, the final bias-aware robust confidence interval is defined as $CI_{1-\alpha}(\hat{\kappa}^*(\hat{\lambda}^*))$ as in (2.9).

3 Asymptotic Size Control

In this section, we show the bias-aware CI proposed in (2.9) has asymptotic size control, which relies on the following assumptions.

Assumption 1. 1. *The graphon $\{Y_{ij}\}_{i\in[N],j\in[M]}$ is generated according to* (2.1) *with $\{X_i, A_i\}_{i=1}^N$, $\{W_j, B_j\}_{j=1}^M$, $\{U_{ij}\}_{i\in[N],j\in[M]}$, $\{v_{ij}\}_{i\in[N],j\in[M]}$ be sequences of i.i.d. random variables that are independent, $M/n \to \phi \in (0,1)$ and $n = M + N$.*

2. *The conditional distribution $\mathbb{P}(Y_{ij} = 1|W_i, X_j)$ is as defined in* (2.2)*, where $\sigma^2 \leq \overline{\sigma}^2$ for some known upper bound $\overline{\sigma}^2$.*

3. *Let $\theta = (\alpha, \beta^\top)^\top \in \mathbb{A} \times \mathbb{B} = \Theta$, the parameter spaces $\mathbb{A}$ and $\mathbb{B}$ are compact.*

4. *Suppose that $\max_{i\in[N],j\in[M]} ||Z_{ij}||_2 = o(n^{1/2})$ with probability one, and*
$$e(R_{ij}^\top \theta_{0,n}) \in \left[\frac{c_1}{n}, \frac{c_2}{n}\right]$$
for some constants $0 < c_1 \leq c_2 < \infty$, where $R_{ij} = (1, Z_{ij}^\top)^\top$.

5. *There exist constants $0 < \lambda_1 < \lambda_2 < \infty$ such that, with probability approaching one,*
$$\lambda_1 \leq \lambda_{\min}\left(\frac{1}{NM}\sum_{i=1}^N\sum_{j=1}^M R_{ij}R_{ij}^\top\right) \leq \lambda_{\max}\left(\frac{1}{NM}\sum_{i=1}^N\sum_{j=1}^M R_{ij}R_{ij}^\top\right) \leq \lambda_2,$$
where for a symmetric matrix A, $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the minimum and maximum eigenvalues of A, respectively.

6. *Let $G(\alpha, \beta) = \mathbb{E}\alpha \exp(Z_{ij}^\top \beta) H_{ij} R_{ij}^\top$ and $G_0 = G(\alpha_0, \beta_0)$. Then, $\lambda_{max}(G(\alpha, \beta) - G_0) \leq C(|\alpha - \alpha_0| + ||\beta - \beta_0||_2)$ for a positive constant $C < \infty$.*

Condition 1 of Assumption 1 describes the graph formation and ensures that $\{Y_{ij}\}_{i\in[N],j\in[M]}$ is X-W exchangeable. Condition 2 restricts the conditional probability to satisfy equation (2.2). Condition 3 is a regularity condition on the parameter space Θ. Condition 4 bounds the largest magnitude of the covariates and the rate at which the probability of forming a link decays. This condition, jointly with 3, controls the rate of sparsity. Condition 5 is an identification condition which ensures that $\theta_{0,n}$ is identified. Finally, Condition 6 represents a Lipschitz continuity, which is needed to estimate G_0 due to the increasing dimension of the sieve space. Conditions 1-5 of Assumption 1 have been used in Graham (2022).

The following assumption provides the smoothness conditions used to provide a tractable approximation to the effect of the local misspecification.

Assumption 2. *1.* $\mathbb{E}\gamma_n(D_{ij}, 0, \theta_{0,n} + t)$ *is differentiable in* t *and*

$$\sup_{t \in \mathcal{N}(0)} ||\partial_t \mathbb{E}\gamma_n(D_{ij}, 0, \theta_{0,n} + t) - \Gamma(t)||_2 \to 0,$$

where $\mathcal{N}(0) \in \Re^{d_z+1}$ *is an arbitrary neighborhood of* 0*. Also, let* $\Gamma(0)$ *be denoted as* Γ*.*

2. There exists $\Delta_v \equiv \frac{1}{2}\lim_{n\to\infty} \mathbb{E}\partial_{vv}\gamma_n(D_{ij}, v, \theta_{0,n})|_{v=0}$ *such that*

$$\sqrt{n}\mathbb{E}\left(\int \gamma_n(D_{ij}, n^{-1/4}v, \theta_{0,n})\pi(v)dv - \gamma_n(D_{ij}, 0, \theta_{0,n})\right) \to \sigma^2\Delta_v. \tag{3.1}$$

Condition 1 of Assumption 2 states that, in the absence of local misspecification, the parameter of interest is Gateaux differentiable at $\theta_{0,n}$. Meanwhile, Condition 2 requires the parameter of interest to be twice-differentiable in the local misspecification component. This condition ensures that the bias in the parameter of interest induced by the local misspecification is non-vanishing at a rate $\sqrt{n}$. All the parameters of interest in Examples 1-6 satisfy this assumption.

Assumption 3. *Suppose* $\|\hat{\kappa}^*(\hat{\lambda}^*) - \kappa^*\|_2 = o_P(k_n^{-1/2})$ *and* $||\kappa^*||_2 \leq C < \infty$ *for some constant* $C > 0$*, where* $\hat{\kappa}^*(\hat{\lambda}^*)$ *and* κ^* *are defined in Section 2.3.*

Assumption 3 provides a rate of approximation for κ^*. This assumption is nontrivial as κ^* implicitly depends on the sample size and its dimension κ_n diverges as the sample size increases. We provide primitive conditions and verify this assumption in the appendix.

Assumption 4. *Let* $\Omega_{ij} = \frac{nVar(S_{Y,ij}|X_i,A_i,W_j,B_j)H_{ij}H_{ij}^\top}{NM}$*. There exists a constant* $c > 0$ *such that* $\lambda_{\min}(\mathbb{E}\Omega_{ij}) \geq c$*. For* $a \in [\underline{a}, 1]$ *for some* $\underline{a} \in (0, 1)$*, we have*

$$\sup_{a\in[\underline{a},1]} \left\| \frac{1}{NM} \sum_{j=1}^{\lceil aM \rceil} \sum_{i=1}^{N} (\Omega_{ij} - \mathbb{E}\Omega_{ij}) \right\|_{op} = O_P(r_n)$$

such that $k_n r_n = o(1)$ *and* $k_n = o(n^{1/2})$*.*

Assumption 4 ensures that the asymptotic variance of the score is well-defined in the limit. Moreover, it provides an upper bound for the deviations of Ω_{ij} from its mean under the uniform operator norm. We verify this assumption in the appendix.

Assumption 5. *The initial estimator satisfies* $\hat{\theta}_{initial} - \theta_0 = O_P(n^{-1/2})$*.*

In the appendix, we show that the logistic MLE satisfies this requirement.

The next theorem characterizes the asymptotic distribution of robust one-step estimator $\hat{\Psi}(\hat{\kappa})$.

Theorem 3.1. *Suppose Assumptions 1-5 hold. Then, the one-step robust estimator $\hat{\Psi}_n(\hat{\kappa})$ has the following asymptotic distribution*

$$\sqrt{n}\left(\hat{\Psi}(\hat{\kappa}) - \Psi_{0,n}\right) \rightsquigarrow \mathcal{N}\left(\sigma^2\left[\Delta_v + \kappa^\top \mathcal{B}_\theta\right], \Sigma_\gamma + 2\Sigma_{H,\gamma}^\top \kappa + \kappa^\top \Sigma_H \kappa\right). \tag{3.2}$$

Theorem 3.1 describes the asymptotic distribution of the one-step robust estimator. There are two main insights that be drawn from Theorem 3.1. First, the estimator $\hat{\Psi}_n$ is asymptotically biased, and the magnitude of this bias is determined by the unknown parameter σ^2, which indexes the degree of local misspecification. Consequently, the asymptotic distribution in Theorem 3.1 cannot be immediately used to conduct inference on $\Psi_{0,n}$. Second, the local misspecification does not have higher-order effects; thus, its complete effect concentrates on the asymptotic bias. In other words, the asymptotic variance does not depend on the local misspecification parameter σ^2.

The next theorem shows that our bias-aware confidence interval introduced in Section 2.1 has the correct asymptotic size control.

Theorem 3.2. *Suppose Assumptions 1–5 hold and the bias-aware CI, i.e., $\widehat{CI}(\hat{\kappa}^*(\hat{\lambda}^*))$, is constructed as* (2.9)*. Then, for the significance level $\alpha \in (0, 1)$, we have*

$$\liminf_{n\to\infty} \mathbb{P}(\Psi_{0,n} \in \widehat{CI}_{1-\alpha}(\hat{\kappa}^*(\hat{\lambda}^*))) \geq 1 - \alpha.$$

This Theorem 3.2 states that the bias-aware confidence interval, which is derived under the robust methodology, yields valid inference on the structural parameter of interest under local misspecification. Notice, that this result holds even when there is no local misspecification. Therefore the bias-aware confidence intervals introduced in Section 2.1 are asymptotically valid for the structural parameter of interest under sparse network asymptotics, both in the correctly specified and locally misspecified case.

4 Asymptotic Optimality

In this section, we show that the one-step robust estimator introduced in 2.1 is minimax optimal in the sense that it attains the smallest confidence interval length within the class of regular estimators.

Consider an arbitrary estimator $\breve{\theta}$ which has the following linear approximation:

$$\sqrt{n}(\breve{\theta} - \theta_{0,n}) = \kappa_F^\top \left[\frac{n^{3/2}}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(Y_{ij} - e(R_{ij}^\top \hat{\theta}_{0,n}))F_{ij}\right] + o_P(1)$$

where $F_{ij} \in \Re^{d_F}$ is d_F dimensional vector consists of transformations of (X_i, W_j), d_F is fixed, and

$\kappa_F \in \Re^{d_F \times d_R}$. We further assume the estimator $\breve{\theta}$ is regular, and thus the following assumption holds.

Assumption 6. *Suppose $G_F^\top \kappa_F = I_{d_R}$ with $G_F = \mathbb{E}\alpha_0 \exp(Z_{ij}^\top \beta_0) F_{ij} R_{ij}^\top$.*

Consider the plug-in estimator $\breve{\Psi}$ of $\Psi_{0,n}$ defined as

$$\breve{\Psi} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \gamma(D_{ij}, 0, \breve{\theta}).$$

Using parallel arguments to those in Section 2.1, we can show that

$$\sqrt{n}(\breve{\Psi} - \Psi_{0,n}) \rightsquigarrow \mathcal{N}(\sigma^2(\Delta_V + \kappa_F^\top B_F), \Sigma_\gamma + 2\Sigma_{F,\gamma}^\top \kappa_F + \kappa_F^\top \Sigma_F \kappa_F),$$

$B_F = \frac{1}{2} \lim_{n\to\infty} \mathbb{E}\left[e\left(R_{ij}^\top \theta_{0,n}\right)\left(e\left(R_{ij}^\top \theta_{0,n}\right) - 1\right)\left(1 - 2e\left(R_{ij}^\top \theta_{0,n}\right)\right) F_{ij}\right]$, and Σ_F and $\Sigma_{F,\gamma}$ are similarly defined as Σ_H and $\Sigma_{H,\gamma}$ with H_{ij} replaced by F_{ij}. Then, the infeasible bias-aware CI based on the plug-in estimator $\breve{\Psi}$ is

$$CI_{1-\alpha}^{\text{plug-in}} \equiv \breve{\Psi} \pm CV_\alpha\left(\frac{\overline{\sigma}^2(\Delta_V + \kappa_F^\top B_F)}{(\Sigma_\gamma + 2\Sigma_{F,\gamma}^\top \kappa_F + \kappa_F^\top \Sigma_F \kappa_F)^{1/2}}\right) \frac{(\Sigma_\gamma + 2\Sigma_{F,\gamma}^\top \kappa_F + \kappa_F^\top \Sigma_F \kappa_F)^{1/2}}{\sqrt{n}}, \tag{4.1}$$

with asymptotic length

$$L^{\text{plug-in}} = 2CV_\alpha\left(\frac{\overline{\sigma}^2(\Delta_V + \kappa_F^\top B_F)}{(\Sigma_\gamma + 2\Sigma_{F,\gamma}^\top \kappa_F + \kappa_F^\top \Sigma_F \kappa_F)^{1/2}}\right)(\Sigma_\gamma + 2\Sigma_{F,\gamma}^\top \kappa_F + \kappa_F^\top \Sigma_F \kappa_F)^{1/2}.$$

Assumption 7. *Suppose there exists a matrix $\Pi_F \in \Re^{d_F \times \kappa_n}$ such that*

$$F_{ij} = \Pi_F H_{ij} + \delta_{ij},$$

and $\mathbb{E}||\delta_{ij}||_2^2 = o(1)$.

This assumptions states that the finite-dimensional vector F_{ij} can be approximated well by the sieve basis H_{ij}.

Theorem 4.1. *Suppose Assumptions 1–7 holds. Then, we have*

$$\hat{L}(\hat{\kappa}^*(\hat{\lambda}^*)) \leq L^{\textit{plug-in}} + o_P(1).$$

Theorem 4.1 states that the one-step robust estimator is minimax optimal within a class of regular estimator. In particular, it shows that the one-step robust estimator attains the shortest confidence interval length with respect to any other regular estimator $\breve{\Psi}$.

5 Application to Specific Cases

In this section, we consider two stylized examples of parameters of interest. In the first case, the parameter of interest is one element of the vector of homophily parameters. In the second case, we focus on the average out-degree parameter. We provide analytic expressions for the sensitivity parameters and characterize the bias-aware confidence intervals.

Case 1: Homophily parameters

Suppose that $\Psi_{0,n} = \beta_0^{(k)}$ for $k = 1, \cdots, K$, then the one-step robust estimator is

$$\hat{\Psi}(\hat{\kappa}) = \iota_k^\top \hat{\beta}_n + \hat{\kappa}^\top \left[\frac{n}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} (Y_{ij} - e(R_{ij}^\top \hat{\theta}_n)) H_{ij} \right],$$

where ι_k is a Kth dimensional vector of zeros with a one in the kth position. The robust estimator has the asymptotic linear expansion

$$\begin{aligned}\sqrt{n}\left(\hat{\Psi}(\hat{\kappa}) - \Psi_{0n}\right) =& \hat{\kappa}^\top \left\{ \frac{n}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \sqrt{n} \left\{ Y_{ij} - \int e(R_{ij}^\top \theta_{0,n} + n^{-1/4} v) \pi(v) dv \right\} H_{ij} \right\} \\ &+ \sigma^2 \hat{\kappa}^\top B_\theta + o_P(1)\end{aligned}$$

where $B_\theta = \frac{1}{2} \lim_{n\to\infty} \mathbb{E}\left[e\left(R_{ij}^\top \theta_{0,n}\right) \left[e\left(R_{ij}^\top \theta_{0,n}\right) - 1 \right] \left[1 - 2e\left(R_{ij}^\top \theta_{0,n}\right) \right] H_{ij} \right]$, which is estimated using

$$\widehat{B}_\theta = \frac{1}{2} \left\{ \frac{n}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \left[\frac{\exp(R_{ij}^\top \hat{\theta}_n)}{\left(1 + \exp(R_{ij}^\top \hat{\theta}_n)\right)^2} \times \frac{1 - \exp(R_{ij}^\top \hat{\theta}_n)}{\left(1 + \exp(R_{ij}^\top \hat{\theta}_n)\right)} \right] H_{ij} \right\}.$$

The optimal value of the sensitivity parameter is determined as the solution to the following optimization

$$\kappa_\lambda^* = \arg\min_{\kappa} \kappa^\top \left\{ \Sigma_H + \lambda \overline{\sigma}^2 B_\theta B_\theta^\top \right\} \kappa \quad \text{s.t.} \quad G^\top \kappa = \iota_k$$

where λ denotes the Lagrange multiplier associated with the worst-bias constraint and $G \equiv \mathbb{E}\left[\alpha_0 \exp\left(Z_{ij}^\top \beta_0\right) H_{ij} R_{ij}^\top \right]$, which is estimated using

$$\widehat{G} = \frac{n}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \left\{ \frac{\exp\left(R_{ij}^\top \hat{\theta}_n\right)}{\left(1 + \exp\left(R_{ij}^\top \hat{\theta}_n\right)\right)^2} \right\} H_{ij} R_{ij}^\top$$

The optimal sensitivity parameter is given by

$$\kappa_\lambda^* = \Upsilon_\lambda^{-1} G \left(G^\top \Upsilon_\lambda^{-1} G \right)^{-1} \iota_k$$

where $\Upsilon_\lambda = (\Sigma_H + \lambda \overline{\sigma}^2 B_\theta B_\theta^\top)$. Finally, given $\alpha \in (0,1)$, the final bias-aware robust confidence interval is defined as

$$\widehat{\mathrm{CI}}_{1-\alpha}(\hat{\Psi}(\hat{\kappa}_{\lambda^*})) = \hat{\Psi}(\hat{\kappa}_{\lambda^*}) \pm CV_\alpha \left(\frac{\overline{\sigma}^2 \kappa_{\lambda^*}^\top \widehat{B}_\theta}{(\hat{\kappa}_{\lambda^*}^\top \hat{\Sigma}_H \hat{\kappa}_{\lambda^*})^{1/2}} \right) \frac{(\hat{\kappa}_{\lambda^*}^\top \hat{\Sigma}_H \hat{\kappa}_{\lambda^*})^{1/2}}{\sqrt{n}}.$$

Case 2: Average out-degree parameter

Let the parameter of interest be given by

$$\Psi_{0,n} = N \mathbb{E} \int e\left[R_{ij}^\top \theta_{0,n} + n^{-1/4} v \right] \pi(v) dv.$$

The one-step robust estimator of $\Psi_{0,n}$ is

$$\hat{\Psi}(\hat{\kappa}) = \frac{1}{M} \sum_{i=1}^{N} \sum_{j=1}^{M} e\left[R_{ij}^\top \hat{\theta}_n \right] + \hat{\kappa}^\top \left[\frac{n}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} (Y_{ij} - e(R_{ij}^\top \hat{\theta}_n)) H_{ij} \right],$$

which has the following asymptotic linear expansion

$$\begin{aligned}
\sqrt{n}\left(\hat{\Psi}(\hat{\kappa}) - \Psi_{0n} \right) =& \frac{\sqrt{n}}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} N \left\{ e(R_{ij}^\top \theta_{0,n}) - \mathbb{E} e(R_{ij}^\top \theta_{0,n}) \right\} \\
&+ \hat{\kappa}^\top \left\{ \frac{\sqrt{n}}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} n \left\{ Y_{ij} - \int e(R_{ij}^\top \theta_{0,n} + n^{-1/4} v) \pi(v) dv \right\} H_{ij} \right\} \\
&+ \sigma^2 \left[\Delta_v + \hat{\kappa}^\top B_\theta \right] + o_P(1)
\end{aligned}$$

where

$$\begin{aligned}
\Delta_v &= \frac{1}{2} \lim_{n \to \infty} N \mathbb{E} \left[e\left(R_{ij}^\top \theta_{0,n} \right) \left[e\left(R_{ij}^\top \theta_{0,n} \right) - 1 \right] \left[1 - 2e\left(R_{ij}^\top \theta_{0,n} \right) \right] \right] \\
B_\theta &= \frac{1}{2} \lim_{n \to \infty} \mathbb{E} \left[e\left(R_{ij}^\top \theta_{0,n} \right) \left[e\left(R_{ij}^\top \theta_{0,n} \right) - 1 \right] \left[1 - 2e\left(R_{ij}^\top \theta_{0,n} \right) \right] H_{ij} \right],
\end{aligned}$$

with estimators given by

$$\widehat{\Delta}_v = \frac{n}{2M}\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\frac{\exp\left(R_{ij}^\top\hat{\theta}_n\right)}{\left(1+\exp\left(R_{ij}^\top\hat{\theta}_n\right)\right)^2}\left\{\frac{1-\exp\left(R_{ij}^\top\hat{\theta}_n\right)}{1+\exp\left(R_{ij}^\top\hat{\theta}_n\right)}\right\}\right]$$

$$\widehat{B}_\theta = \frac{n}{2NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\frac{\exp\left(R_{ij}^\top\hat{\theta}_n\right)}{\left(1+\exp\left(R_{ij}^\top\hat{\theta}_n\right)\right)^2}\left\{\frac{1-\exp\left(R_{ij}^\top\hat{\theta}_n\right)}{1+\exp\left(R_{ij}^\top\hat{\theta}_n\right)}\right\}\right] H_{ij}.$$

The optimal value of the sensitivity parameter is determined as the solution to the following optimization

$$\begin{aligned}\kappa_\lambda^* =& \arg\min_\kappa \left(\Sigma_\gamma + \lambda\overline{\sigma}^2\Delta_v^2\right) + 2\left[\Sigma_{H,\gamma} + \lambda\overline{\sigma}^2\Delta_{\gamma_2}B_\theta^\top\right]\kappa + \kappa^\top\left\{\Sigma_H + \lambda\overline{\sigma}^2\bar{B}_\theta\bar{B}_\theta^\top\right\}\kappa \\ &\text{s.t.}\quad G^\top\kappa = \Gamma\end{aligned}$$

where λ denotes the Lagrange multiplier associated with the worst-bias constraint and $G = \mathbb{E}\left[\alpha_0\exp\left(Z_{ij}^\top\beta_0\right)H_{ij}R_{ij}^\top\right]$ and $\Gamma = \lim_{n\to\infty} N\mathbb{E}\left[e\left(R_{ij}^\top\theta_{0,n}\right)\left\{1-e\left(R_{ij}^\top\theta_{0,n}\right)\right\}R_{ij}\right]$. The corresponding estimators are

$$\widehat{G} = \frac{n}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left\{\frac{\exp\left(R_{ij}^\top\hat{\theta}_n\right)}{\left(1+\exp\left(R_{ij}^\top\hat{\theta}_n\right)\right)^2}\right\}H_{ij}R_{ij}^\top$$

$$\widehat{\Gamma} = \frac{1}{M}\sum_{i=1}^{N}\sum_{j=1}^{M}\left\{\frac{\exp\left(R_{ij}^\top\hat{\theta}_n\right)}{\left(1+\exp\left(R_{ij}^\top\hat{\theta}_n\right)\right)^2}\right\}R_{ij}.$$

The optimal value of the sensitivity parameter is

$$\kappa_\lambda^* = -\Upsilon_\lambda^{-1}\left\{A_\lambda^\top - G\left(G^\top\Upsilon_\lambda^{-1}G^\top\right)^{-1}\left[\Gamma + G^\top\Upsilon_\lambda^{-1}A_\lambda^\top\right]\right\}$$

where $\Upsilon_\lambda = (\Sigma_H + \lambda\overline{\sigma}^2B_\theta B_\theta^\top)$ and $A_\lambda = \left[\Sigma_{H,\gamma} + \lambda\overline{\sigma}^2\Delta_v B_\theta^\top\right]$. Finally, given $\alpha\in(0,1)$, the final bias-aware robust confidence interval is

$$\widehat{\mathrm{CI}}_{1-\alpha}(\hat{\Psi}(\hat{\kappa}_{\lambda^*})) = \hat{\Psi}(\hat{\kappa}_{\lambda^*}) \pm CV_\alpha\left(\frac{\overline{\sigma}^2|\hat{\Delta}_v + \hat{\kappa}_{\lambda^*}^\top\hat{B}_\theta|}{(\hat{\Sigma}_\gamma + 2\hat{\Sigma}_{H,\gamma}^\top\hat{\kappa}_{\lambda^*} + \hat{\kappa}_{\lambda^*}^\top\hat{\Sigma}_H\hat{\kappa}_{\lambda^*})^{1/2}}\right)\frac{(\hat{\Sigma}_\gamma + 2\hat{\Sigma}_{H,\gamma}^\top\hat{\kappa}_{\lambda^*} + \hat{\kappa}_{\lambda^*}^\top\hat{\Sigma}_H\hat{\kappa}_{\lambda^*})^{1/2}}{\sqrt{n}}.$$

6 Monte Carlo Simulations

This section presents simulation evidence for the finite sample performance of the one-step robust estimator introduced in Section 3. We consider a wide array of Monte Carlo designs that are meant to capture differences in the specification of the local misspecification, sample sizes and sparsity of the network.

The bipartite network is modelled according to the following data-generating process. For any $i \in [N]$ and $j \in [M]$, the observed attributes X_i and W_j are drawn from independent and identically distributed log-normal distributions with mean $-1/4$ and variance $1/2$. The dyad-specific attributes are computed to account for assortative matching. In particular, we define $Z_{ij} = \log(X_i \cdot W_j)$, which will be distributed normally with mean equal to $-1/2$ and variance 1. The sieves basis is computed using a Hermite polynomial approximation of Z_{ij} of order k_n. As an alternative sieve basis, we consider the tensor product of polynomial expansions on X_i and W_j, which yields similar qualitative results. We also consider sieve basis of different orders with similar qualitative results.

The unobserved heterogeneity A_i and B_j are drawn from independent and identically distributed log-normal distributions with mean $-1/12$ and standard deviation given by $1/\sqrt{6}$. Finally, U_{ij} is distributed as a standard exponential distribution. The bipartite graphon is simulated according to the equation

$$Y_{ij} = 1\left[\alpha_0 - \log(n) + Z_{ij}^{\top}\beta_0 + \log(A_i) + \log(B_i) + n^{-1/4}v_{ij} + U_{ij} \geq 0\right], \tag{6.1}$$

where the term $\log(n)$ will ensure that the average degree of this network is bounded. In particular, under the current setting, the probability of establishing a link decreases at a rate n^{-1}.

The distribution of the misspecification component v_{ij} is simulated according to four different designs: (i) $v_{ij} \sim N(0, \sigma)$, (ii) $v_{ij} \sim \text{Logistic}(0, \sigma)$, (iii) $v_{ij} \sim \text{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$, and (iv) $v_{ij} \sim \text{Laplace}(0, \sigma)$ for $\sigma^2 \in \{0, 1, 2, 3, 4\}$. All designs ensure v_{ij} has mean zero and variance proportional to σ^2. A $\sigma^2 = 0$ represents the absence of local misspecification. The upper bound on the magnitude of local misspecification σ^2 is fixed as $M = 4$ across all designs. That is, it corresponds with the largest value taken by σ^2. The true DGP design is completed by setting $\alpha_0 = \log(2.56)$, $\beta_0 = \log(4)$, and network size equal to $N, M = 150, 200, 250, 300$ with $n = N + M$.

The one-step robust estimator is computed using the definition in equation (2.4). As an initial estimator $\hat{\theta}_{\text{initial}}$, we consider a Logistic regression. We compare the performance of the one-step robust estimator with that of a plug-in estimator with two different initial estimators: (i) a Logistic regression and (ii) a Poisson regression. Given the exponential distribution of U_{ij}, the logistic estimator $\hat{\theta}_n$ will present a bias of order $o(n^{-2})$, as discussed in Graham (2022). This bias is negligible concerning the effect generated by the local misspecification, which induces a nonvanishing bias in the asymptotic distribution of order $n^{-1/2}$. On the contrary, the initial Poisson estimator corresponds to that of a Composite Maximum Likelihood Estimator when the distribution of the error term is known. We report bias-aware confidence intervals for all the estimators, the one-step robust estimator and both plug-in estimators.

For the parameter of interest $\Psi_{0,n}$, we consider two cases: (i) $\Psi_{0,n}$ is equal to the homophily parameter β_0, and (ii) $\Psi_{0,n}$ is the average out-degree of the network $N\mathbb{E}\int \exp(R_{ij}^\top \theta_{0,n}+n^{-1/4}v)\pi(v)dv$. Notice that in case (i), the parameter of interest is fixed across all the designs of σ^2. Meanwhile, in the second setting, the parameter of interest is affected directly by the local misspecification.

6.1 Homophily parameters

Table 4 summarizes results from computing the one-step robust estimator across 1,000 Monte Carlo replications when the parameter of interest is $\Psi_{0,n} = \beta_0$ and $v_{ij} \sim \mathcal{N}(0, \sigma^2)$. The table includes the true value of β_0, along with the mean coefficient estimates (coeff.), $\sqrt{n}$-bias, standard error (s.e.), standard error-to-standard deviation ratio (s.e./sd), root-mean-squared error (RMSE), 95% confidence interval (conf. int.), confidence interval width (width), and 95% coverage probabilities (95% CP). The final column shows the average degree of the network. These values were calculated over all simulations.

The top panel in Table 4 shows the results of computing the robust estimator in a network of size $n = 300$ and across misspecification designs $\sigma^2 \in \{0, 1, 2, 3, 4\}$. The mean estimate and $\sqrt{n}$-Bias show that the robust estimator approximates well the true DGP value. Moreover, the standard error-to-standard deviation ratio (s.e./sd) indicates that the estimator's sampling variability is well approximated. Finally, the bias-aware confidence intervals have the correct asymptotic coverage under sparse asymptotic designs.

The bottom panels of Table 4 report the results of the robust estimator in larger networks with sizes equal to $n = 400, 500$ and 600. As the sample size grows, the point estimates approximate more precisely the true parameter value. Note that the $\sqrt{n}$-Bias does not vanish completely. This is expected as it results from scaling up the average bias by a $\sqrt{n}$ factor. This component captures the first-order effect that the local misspecification has on the asymptotic distribution of the robust estimator, which is non-vanishing. Moreover, the width of the bias-aware confidence intervals decreases uniformly as n grows throughout all the misspecification designs and even when the average network degree is as low as 0.5%.

Tables 5 and 6 report the results of the plug-in estimators with an initial logistic and Poisson regression, respectively. Relative to the plug-in estimator with an initial logistic estimator, it is clear that the robust estimator exhibits smaller RMSEs and shorter 95% bias-aware confidence intervals. In fact, the width of the confidence intervals obtained under the robust methodology is, on average, 2.5% smaller when $n = 300$. The largest correction happens when $\sigma^2 = 4$, with a confidence interval that is 2.58% shorter. Even as both estimators become more precise in larger networks, the robust estimator attains shorter confidence intervals. They are, on average, 1.8% shorter when $n = 400$, 1.45% shorter when $n = 500$, and 1.20% when $n = 600$.

The robust estimator also outperforms the plug-in estimator with an initial Poisson regression in Table 6. When comparing the two, the robust estimator exhibits smaller RMSEs and shorter 95% bias-aware confidence intervals. Albeit, the margins are smaller as the plug-in estimator coincides with the composite MLE. We observe that the average length reductions are 1.17% for $n = 300$, 0.87% for $n = 400$, 0.71% for $n = 500$, and 0.59% for $n = 600$. Notably, the most substantial improvements happen when $\sigma^2 = 4$, that is, the magnitude of the local misspecification coincides with its upper bound (M). During these instances, the confidence intervals under robust inference are 1.25% smaller for $n = 300$, 0.92% smaller for $n = 400$, 0.75% smaller for $n = 500$, and 0.62% smaller for $n = 600$.

Appendix Tables 10 - 18 report the simulation results when v_{ij} is distributed (i) Logistic$(0, \sigma)$, (ii) U$\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$, and (iii) Laplace$(0, \sigma)$. Qualitatively, we observe similar patterns. This evidence suggests that there is an advantage in terms of efficiency of computing bias-aware confidence intervals relative to the initial estimator, even when the parameter of interest is relatively simple.

6.2 Average out-degree

Table 7 summarizes results from computing the one-step robust estimator across 1,000 Monte Carlo replications when the parameter of interest is the average out-degree of the network, $\Psi_{0,n} = N\mathbb{E}\int \exp\left[R_{ij}^{\top}\theta_{0,n} + n^{-1/4}v\right]\pi(v)dv$, and $v_{ij} \sim \mathcal{N}(0, \sigma^2)$. Notice that, in this case, the local misspecification directly affects the parameter of interest, and thus, the true value of $\Psi_{0,n}$ varies across different specifications of σ^2.

The results in Table 7 suggest that the robust estimator yields reliable inference for the parameter of interest, and its performance improves in larger networks, notwithstanding the large degree of sparsity. The $\sqrt{n}$-Bias is relatively small and shows a decaying pattern as n grows. The bias-aware confidence intervals have the correct asymptotic coverage, and their width decreases uniformly as n grows and throughout all the misspecification designs.

Tables 8 and 9 report the results of the plug-in estimators with an initial logistic and Poisson regression, respectively. When comparing the plug-in estimator with an initial logistic regression and the robust estimator, it is clear that the robust estimator improves significantly on the plug-in estimator. Not only it exhibits smaller RMSEs, but it also offers a larger correction on the 95% bias-aware confidence intervals. In particular, the average length reductions are 12.60% for $n = 300$, 10.41% for $n = 400$, 9.60% for $n = 500$, and 8.17% for $n = 600$. The same pattern is observed when comparing the robust estimator with the plug-in estimator with an initial Poisson regression in Table 6. Although this second plug-in estimator provides better asymptotic guarantees, the results show that using the robust estimator might yield an average reduction of the confidence interval's width of 3.12% for $n = 300$, 2.40% for $n = 400$, 2.78% for $n = 500$, and 2.11% for $n = 600$.

Appendix Tables 19 - 27 report the simulation results when v_{ij} is distributed (i) Logistic $(0, \sigma)$, (ii) U $\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$, and (iii) Laplace $(0, \sigma)$. Qualitatively, we observe similar patterns. This evidence suggests that computing the optimal confidence intervals of the robust estimator might lead to a significant improvement when the parameter of interest is affected directly by the local misspecification.

7 Empirical Application

In this section, we implement the methodology developed in Section 3 to study a network of scientific collaborations among economists. We utilize articles published in leading American journals of general interest from 2000 to 2006 to construct a bipartite network connecting authors and articles. The goal is to estimate the factors influencing an author's decision to participate in a paper, thereby establishing a scientific collaboration with other authors who may be involved in that project.

This study contributes to the growing literature on scientific collaborations in economics. Most of the existing literature has focused on describing stylized features of this network (Newman 2001; Goyal et al. 2006 and Boschini and Sjögren 2007) or measuring the effect that these collaborations have on research productivity (Ductor, Fafchamps, Goyal, and Van der Leij 2014; Ductor 2015 and Colussi 2018). Fewer studies have analyzed the mechanisms that drive the formation of these scientific collaborations (Fafchamps, Van der Leij, and Goyal 2010; Anderson and Richards-Shubik 2022 and Hsieh et al. 2022).[2] We extend these studies by considering a larger number of factors that can influence an author's decision to collaborate on a project while using a computationally tractable method that accounts for two-sided heterogeneity. More importantly, this is the first paper to study the effects that local misspecification has on a bipartite network of scientific collaboration and conduct robust inference on the parameters of interest.

The data includes all papers published in top American economic journals from 2000 to 2006: the American Economic Review, Econometrica, Journal of Political Economy, and Quarterly Journal of Economics. It contains author's information, such as gender, PhD university, graduation year, research fields, institution of employment and position. For articles, the dataset includes citation count since the year of publication, authors' names, publication issue, number of pages, references, three-digit JEL code, and keywords.[3] We used this data to define a single, static bipartite network of collaborations that includes a total of 1870 authors and 1600 articles.[4]

We examine the formation probability of a collaboration link using the equation (2.2) and compare the one-step robust estimator's performance with plug-in estimators based on initial logistic and Poisson regressions. The parameters of interest include assortative matching parameters and

[2] Bonhomme (2021) introduces a structural approach to identify individual contributions in teams production.

[3] We thank Tommaso Colussi for sharing this data with us.

[4] Our final sample includes only unique observations of matched authors and articles.

the network's average out-degree. We control for author and project-specific characteristics, along with author-project pair-specific observed attributes. In terms of individual author's attributes, we include binary variables for the categories of female gender and junior economist position and control for the average number of citations received by the author.[5,6] As project-specific attributes, we include the number of authors in the project, indices for the share of female and senior authors in the project, and binary variables indicating whether some authors share the same institution or obtained their PhD from the same university.[7]

As main mechanisms for assortative matching, we consider: (i) similarity between author i' research fields and project j's classification topic (`jelcode`), (ii) sorting of more productive authors into projects with higher impact (`citations`), (iii) gender sorting across collaborations (`gender`), and (iv) collaboration between junior and senior scientists (`junior_senior`). Table 1 collects the estimation results using the robust methodology with an initial Logistic estimator. It reports the coefficient estimates, standard errors, confidence intervals, confidence intervals' width and the adjustment term of the one-step robust estimator.

The results in Table 1 indicate that similarity in the research field of expertise is a strong and positive factor influencing the decision to collaborate on a research project. This result aligns with the findings in Ductor 2015 and Hsieh et al. 2022. Additionally, highly productive authors tend to participate in projects with higher impact with a positive probability. This outcome is consistent with the role of "star" economists discussed in Goyal et al. (2006). The evidence also suggests that, during the sample period, sorting of female authors is a strong positive predictor for the establishment of a collaboration network. This pattern is also observed by Boschini and Sjögren (2007) during the 1991-2002 time period. On the contrary, this analysis suggests that junior scientists are less likely to participate in a project with a large share of senior authors. One likely explanation is that successful collaborations tend to perdure, and it is costly to establish new connections (cf. Hsieh et al. 2022). Notice that the estimates predict a negative constant coefficient, which is expected from a sparse bipartite network. the analysis indicates that there are approximately 2 scientists who collaborate on each project.

When comparing these results to those of the plug-in estimator in Tables 2 and 3, we observe the same patterns in the prediction mechanisms. However, the confidence intervals for the robust estimator are noticeably narrower. Specifically, substantial improvements are evident in predicting the effects of research field similarity, citation counts, the constant coefficient, as well as the average out-degree parameter.

[5] The category of junior economists includes assistant professors and economists in government or research institutions.

[6] The average citation count serves as a proxy measure for the author's productivity.

[7] The category of senior economists includes associate and full professors and senior economists in government or research institutions.

Table 1: Robust Estimator Results

	coeff.	s.e.	[conf. int.]	width	adjm
			Homophily parameters		
jelcode	1.492	0.028	[1.432, 1.552]	0.121	-0.017
citations	0.115	0.004	[0.107, 0.124]	0.017	0.003
gender	5.449	0.163	[5.129, 5.770]	0.641	0.539
junior_senior	-3.297	0.128	[−3.549, −3.046]	0.503	-0.284
constant	-0.833	0.027	[−0.888, −0.777]	0.111	-0.011
			Average out-degree		
$\hat{\Psi}_n$	1.874	0.035	[1.807, 1.942]	0.135	-0.083

[1] Total sample includes $N = 1870$ authors and $M = 1600$ articles.

Table 2: Logit Estimator Results

	coeff.	s.e.	[conf. int.]	width	adjm
			Homophily parameters		
jelcode	1.509	0.041	[1.430, 1.589]	0.159	0.0
citations	0.113	0.008	[0.098, 0.128]	0.030	0.0
gender	4.910	0.163	[4.590, 5.230]	0.641	0.0
junior_senior	-3.014	0.130	[−3.268, −2.760]	0.508	0.0
constant	-0.822	0.076	[−1.004, −0.640]	0.365	0.0
			Average out-degree		
$\hat{\Psi}_n$	1.957	0.112	[1.737, 2.177]	0.441	0.0

[1] Total sample includes $N = 1870$ authors and $M = 1600$ articles.

Table 3: Poisson Estimator Results

	coeff.	s.e.	[conf. int.]	width	adjm
			Homophily parameters		
jelcode	1.510	0.041	[1.430, 1.589]	0.159	0.0
citations	0.089	0.006	[0.077, 0.102]	0.025	0.0
gender	5.410	0.178	[5.047, 5.774]	0.727	0.0
junior_senior	-2.986	0.130	[−3.241, −2.731]	0.510	0.0
constant	-0.820	0.076	[−1.003, −0.638]	0.365	0.0
			Average out-degree		
$\hat{\Psi}_n$	1.874	0.038	[1.800, 1.949]	0.150	0.0

[1] Total sample includes $N = 1870$ authors and $M = 1600$ articles.

8 Conclusion

We study the effects of local misspecification on a bipartite network. We focus on a class of dyadic network models characterized by conditional moment restrictions that are locally misspecified. The magnitude of misspecification is indexed by the sample size and vanishes at a rate $n^{-1/4}$. We utilize this local asymptotic approach to construct a robust estimator that attains the shortest confidence interval length within a prespecified neighborhood of misspecification. Additionally, we introduce bias-aware confidence intervals that account for the effect of the local misspecification. These confidence intervals are asymptotically valid for the structural parameter of interest under sparse network asymptotics, both in the correctly-specified and locally-misspecified case. In an empirical application, we study the formation of a scientific collaboration network among economists. Our analysis documents that homophily in the research field of expertise, sorting of highly productive scientists into higher-impact projects, and participation of female authors in teams with a higher share of female authors are all strong statistical factors that explain the formation of a collaboration network.

A Figures & Tables

Figure 1: Confidence Intervals: β_0

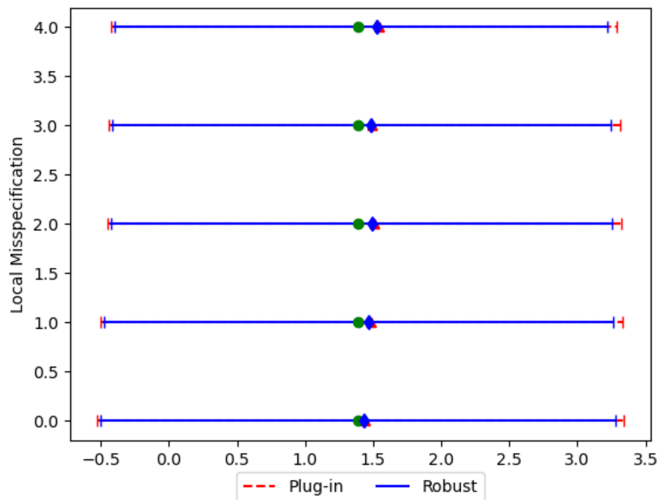

Figure 2: Confidence Intervals: $\Psi_{0,n}$

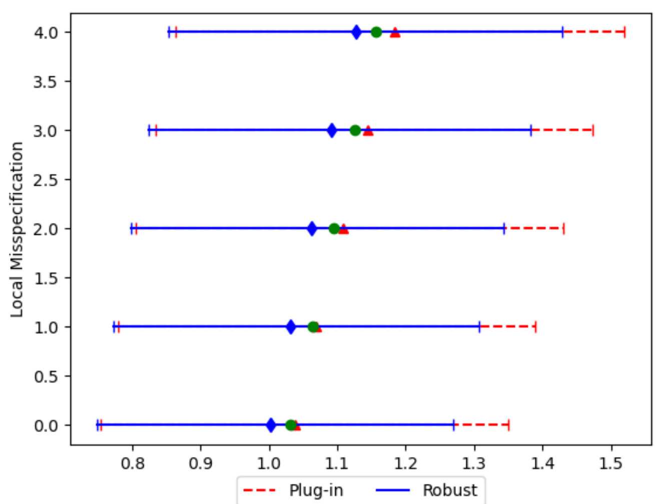

Table 4: $\hat{\beta}_n$ Robust Estimator

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.429	0.738	0.955	0.963	16.968	[−0.503, 3.284]	3.786	0.932	0.010
1.0	1.386	1.463	1.329	0.935	0.948	16.794	[−0.476, 3.267]	3.743	0.940	0.010
2.0	1.386	1.489	1.778	0.925	0.940	16.491	[−0.422, 3.257]	3.679	0.944	0.010
3.0	1.386	1.479	1.611	0.920	0.946	16.403	[−0.414, 3.247]	3.661	0.936	0.010
4.0	1.386	1.524	2.393	0.909	0.936	16.216	[−0.396, 3.224]	3.619	0.940	0.011
					$n = 400$					
0.0	1.386	1.330	−1.134	0.835	0.917	17.098	[−0.316, 3.000]	3.316	0.914	0.007
1.0	1.386	1.366	−0.408	0.826	0.915	16.915	[−0.293, 2.987]	3.280	0.930	0.007
2.0	1.386	1.369	−0.349	0.814	0.931	16.718	[−0.262, 2.981]	3.243	0.932	0.008
3.0	1.386	1.359	−0.537	0.808	0.920	16.590	[−0.248, 2.971]	3.219	0.932	0.008
4.0	1.386	1.378	−0.166	0.803	0.933	16.414	[−0.221, 2.964]	3.185	0.932	0.008
					$n = 500$					
0.0	1.386	1.363	−0.512	0.761	0.922	17.418	[−0.176, 2.852]	3.028	0.930	0.006
1.0	1.386	1.381	−0.128	0.763	0.940	17.285	[−0.177, 2.829]	3.006	0.932	0.006
2.0	1.386	1.347	−0.889	0.756	0.958	17.144	[−0.156, 2.825]	2.980	0.932	0.006
3.0	1.386	1.355	−0.698	0.751	0.957	17.001	[−0.138, 2.817]	2.955	0.930	0.006
4.0	1.386	1.359	−0.601	0.739	0.936	16.820	[−0.115, 2.808]	2.923	0.932	0.006
					$n = 600$					
0.0	1.386	1.432	1.125	0.693	0.963	17.258	[−0.003, 2.738]	2.741	0.932	0.005
1.0	1.386	1.416	0.733	0.687	0.971	17.172	[−0.001, 2.726]	2.727	0.932	0.005
2.0	1.386	1.399	0.321	0.683	0.964	17.021	[0.008, 2.711]	2.703	0.942	0.005
3.0	1.386	1.373	−0.322	0.680	0.948	16.896	[0.017, 2.702]	2.685	0.938	0.005
4.0	1.386	1.379	−0.174	0.674	0.949	16.744	[0.031, 2.691]	2.660	0.938	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Local misspecification v_{ij} is drawn from a $N(0, \sigma)$. [3] Largest degree of misspecification $M = 4$. [4] Sieves dimension $k_n = 2$.

Table 5: $\hat{\beta}_n$ Logistic Estimator

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.445	1.010	0.978	0.975	17.365	[−0.532, 3.344]	3.875	0.932	0.010
1.0	1.386	1.480	1.618	0.956	0.957	17.198	[−0.505, 3.328]	3.833	0.944	0.010
2.0	1.386	1.506	2.074	0.948	0.952	16.902	[−0.451, 3.320]	3.771	0.944	0.010
3.0	1.386	1.496	1.895	0.945	0.959	16.824	[−0.443, 3.312]	3.755	0.936	0.010
4.0	1.386	1.542	2.698	0.933	0.948	16.643	[−0.426, 3.289]	3.715	0.942	0.011
					$n = 400$					
0.0	1.386	1.340	−0.916	0.851	0.926	17.395	[−0.334, 3.040]	3.374	0.914	0.007
1.0	1.386	1.376	−0.205	0.841	0.923	17.217	[−0.311, 3.028]	3.339	0.930	0.007
2.0	1.386	1.380	−0.135	0.828	0.939	17.024	[−0.280, 3.023]	3.303	0.932	0.008
3.0	1.386	1.371	−0.305	0.823	0.929	16.902	[−0.266, 3.013]	3.280	0.936	0.008
4.0	1.386	1.391	0.104	0.819	0.943	16.731	[−0.239, 3.007]	3.247	0.936	0.008
					$n = 500$					
0.0	1.386	1.371	−0.332	0.772	0.929	17.663	[−0.189, 2.882]	3.071	0.934	0.006
1.0	1.386	1.389	0.068	0.773	0.945	17.534	[−0.189, 2.860]	3.049	0.934	0.006
2.0	1.386	1.356	−0.689	0.767	0.964	17.396	[−0.168, 2.856]	3.024	0.932	0.006
3.0	1.386	1.364	−0.493	0.762	0.964	17.257	[−0.151, 2.849]	3.000	0.932	0.006
4.0	1.386	1.368	−0.399	0.751	0.944	17.080	[−0.128, 2.840]	2.969	0.934	0.006
					$n = 600$					
0.0	1.386	1.440	1.305	0.701	0.968	17.461	[−0.012, 2.761]	2.773	0.934	0.005
1.0	1.386	1.424	0.924	0.695	0.976	17.377	[−0.009, 2.750]	2.759	0.934	0.005
2.0	1.386	1.407	0.499	0.691	0.970	17.228	[−0.001, 2.735]	2.736	0.940	0.005
3.0	1.386	1.381	−0.136	0.689	0.954	17.106	[0.008, 2.726]	2.718	0.938	0.005
4.0	1.386	1.387	0.012	0.682	0.955	16.957	[0.022, 2.716]	2.694	0.940	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Local misspecification v_{ij} is drawn from a $N(0, \sigma)$. [3] Largest degree of misspecification $M = 4$. [4] Sieves dimension $k_n = 2$.

Table 6: $\hat{\beta}_n$ Poisson Estimator

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.429	0.735	0.966	0.974	17.159	[−0.524, 3.305]	3.829	0.934	0.010
1.0	1.386	1.463	1.326	0.945	0.957	16.987	[−0.498, 3.288]	3.786	0.944	0.010
2.0	1.386	1.489	1.775	0.936	0.951	16.688	[−0.444, 3.279]	3.723	0.944	0.010
3.0	1.386	1.479	1.608	0.932	0.958	16.605	[−0.436, 3.270]	3.706	0.938	0.010
4.0	1.386	1.524	2.389	0.920	0.947	16.420	[−0.419, 3.246]	3.665	0.942	0.011
					$n = 400$					
0.0	1.386	1.330	−1.136	0.842	0.925	17.241	[−0.330, 3.013]	3.344	0.914	0.007
1.0	1.386	1.366	−0.409	0.833	0.923	17.060	[−0.308, 3.001]	3.308	0.934	0.007
2.0	1.386	1.369	−0.351	0.820	0.939	16.866	[−0.277, 2.995]	3.272	0.932	0.008
3.0	1.386	1.359	−0.539	0.815	0.929	16.740	[−0.263, 2.985]	3.248	0.934	0.008
4.0	1.386	1.378	−0.169	0.810	0.942	16.567	[−0.236, 2.979]	3.214	0.938	0.008
					$n = 500$					
0.0	1.386	1.363	−0.513	0.767	0.929	17.537	[−0.187, 2.862]	3.049	0.936	0.006
1.0	1.386	1.381	−0.129	0.768	0.946	17.407	[−0.187, 2.840]	3.027	0.932	0.006
2.0	1.386	1.346	−0.890	0.762	0.965	17.267	[−0.166, 2.835]	3.002	0.932	0.006
3.0	1.386	1.355	−0.699	0.757	0.965	17.126	[−0.149, 2.828]	2.977	0.936	0.006
4.0	1.386	1.359	−0.602	0.745	0.944	16.947	[−0.127, 2.819]	2.945	0.934	0.006
					$n = 600$					
0.0	1.386	1.432	1.124	0.697	0.968	17.357	[−0.011, 2.745]	2.757	0.932	0.005
1.0	1.386	1.416	0.732	0.691	0.976	17.272	[−0.009, 2.734]	2.743	0.934	0.005
2.0	1.386	1.399	0.320	0.687	0.970	17.122	[0.000, 2.719]	2.719	0.942	0.005
3.0	1.386	1.373	−0.323	0.684	0.954	16.999	[0.009, 2.710]	2.701	0.938	0.005
4.0	1.386	1.379	−0.175	0.678	0.955	16.848	[0.023, 2.699]	2.677	0.940	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Local misspecification v_{ij} is drawn from a $N(0, \sigma)$. [3] Largest degree of misspecification $M = 4$. [4] Sieves dimension $k_n = 2$.

Table 7: $\hat{\Psi}_n$ Robust Estimator

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.032	1.003	−0.506	0.129	0.898	2.333	[0.748, 1.270]	0.521	0.914	0.007
1.0	1.065	1.032	−0.572	0.132	0.885	2.390	[0.773, 1.307]	0.533	0.912	0.007
2.0	1.095	1.063	−0.551	0.135	0.888	2.443	[0.798, 1.343]	0.545	0.908	0.007
3.0	1.126	1.092	−0.580	0.138	0.883	2.500	[0.825, 1.383]	0.558	0.918	0.008
4.0	1.157	1.127	−0.523	0.142	0.886	2.573	[0.854, 1.428]	0.574	0.932	0.008
					$n = 400$					
0.0	1.029	1.002	−0.522	0.114	0.980	2.373	[0.781, 1.241]	0.460	0.936	0.005
1.0	1.053	1.024	−0.571	0.118	0.975	2.419	[0.799, 1.268]	0.469	0.934	0.005
2.0	1.080	1.051	−0.587	0.120	0.973	2.471	[0.820, 1.299]	0.479	0.934	0.005
3.0	1.107	1.077	−0.598	0.123	0.985	2.525	[0.843, 1.333]	0.489	0.934	0.006
4.0	1.135	1.104	−0.612	0.125	0.977	2.580	[0.867, 1.367]	0.500	0.938	0.006
					$n = 500$					
0.0	1.031	1.015	−0.376	0.104	0.929	2.370	[0.812, 1.224]	0.412	0.918	0.004
1.0	1.055	1.036	−0.416	0.106	0.937	2.414	[0.829, 1.249]	0.420	0.920	0.004
2.0	1.078	1.057	−0.470	0.108	0.935	2.465	[0.849, 1.278]	0.429	0.924	0.004
3.0	1.105	1.085	−0.448	0.110	0.943	2.511	[0.869, 1.306]	0.436	0.914	0.004
4.0	1.130	1.105	−0.554	0.111	0.942	2.555	[0.889, 1.333]	0.444	0.922	0.005
					$n = 600$					
0.0	1.034	1.024	−0.250	0.095	0.998	2.369	[0.834, 1.211]	0.377	0.944	0.003
1.0	1.056	1.040	−0.393	0.096	0.978	2.404	[0.852, 1.234]	0.383	0.934	0.004
2.0	1.077	1.066	−0.271	0.098	0.976	2.447	[0.870, 1.259]	0.389	0.938	0.004
3.0	1.099	1.086	−0.336	0.100	0.971	2.488	[0.889, 1.285]	0.396	0.934	0.004
4.0	1.121	1.107	−0.364	0.101	0.972	2.530	[0.908, 1.310]	0.402	0.938	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Local misspecification v_{ij} is drawn from a $N(0, \sigma)$. [3] Largest degree of misspecification $M = 4$. [4] Sieves dimension $k_n = 3$.

Table 8: $\hat{\Psi}_n$ Logistic Estimator

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.032	1.038	0.099	0.140	0.942	2.518	[0.754, 1.349]	0.596	0.952	0.007
1.0	1.065	1.070	0.093	0.142	0.912	2.582	[0.780, 1.389]	0.609	0.956	0.007
2.0	1.095	1.109	0.242	0.146	0.921	2.652	[0.806, 1.431]	0.625	0.956	0.007
3.0	1.126	1.144	0.322	0.150	0.921	2.715	[0.834, 1.473]	0.639	0.958	0.008
4.0	1.157	1.183	0.448	0.154	0.914	2.796	[0.863, 1.520]	0.657	0.960	0.008
					$n = 400$					
0.0	1.029	1.036	0.150	0.121	1.001	2.512	[0.787, 1.301]	0.513	0.968	0.005
1.0	1.053	1.059	0.122	0.124	1.001	2.560	[0.808, 1.331]	0.523	0.968	0.005
2.0	1.080	1.086	0.109	0.128	1.008	2.621	[0.830, 1.364]	0.535	0.968	0.005
3.0	1.107	1.120	0.268	0.131	1.007	2.680	[0.853, 1.399]	0.546	0.970	0.006
4.0	1.135	1.145	0.210	0.133	1.004	2.744	[0.877, 1.435]	0.559	0.978	0.006
					$n = 500$					
0.0	1.031	1.043	0.260	0.109	0.945	2.502	[0.819, 1.275]	0.456	0.956	0.004
1.0	1.055	1.063	0.176	0.111	0.959	2.554	[0.838, 1.303]	0.465	0.948	0.004
2.0	1.078	1.082	0.100	0.114	0.962	2.606	[0.859, 1.333]	0.474	0.954	0.004
3.0	1.105	1.109	0.098	0.116	0.966	2.653	[0.879, 1.361]	0.482	0.952	0.004
4.0	1.130	1.137	0.167	0.118	0.966	2.703	[0.899, 1.390]	0.491	0.948	0.005
					$n = 600$					
0.0	1.034	1.045	0.267	0.100	1.032	2.468	[0.841, 1.251]	0.411	0.960	0.003
1.0	1.056	1.062	0.145	0.101	1.005	2.508	[0.858, 1.275]	0.417	0.952	0.004
2.0	1.077	1.085	0.200	0.103	1.002	2.553	[0.877, 1.301]	0.424	0.950	0.004
3.0	1.099	1.109	0.238	0.105	0.996	2.594	[0.896, 1.326]	0.431	0.952	0.004
4.0	1.121	1.130	0.216	0.107	1.003	2.640	[0.915, 1.353]	0.438	0.956	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Local misspecification v_{ij} is drawn from a $N(0, \sigma)$. [3] Largest degree of misspecification $M = 4$. [4] Sieves dimension $k_n = 3$.

Table 9: $\hat{\Psi}_n$ Poisson Estimator

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.032	1.020	−0.214	0.134	0.951	2.408	[0.763, 1.301]	0.538	0.938	0.007
1.0	1.065	1.053	−0.194	0.137	0.925	2.464	[0.789, 1.339]	0.550	0.938	0.007
2.0	1.095	1.087	−0.138	0.141	0.933	2.525	[0.815, 1.378]	0.564	0.936	0.007
3.0	1.126	1.127	0.018	0.144	0.935	2.582	[0.842, 1.418]	0.577	0.944	0.008
4.0	1.157	1.160	0.048	0.148	0.927	2.654	[0.870, 1.463]	0.593	0.946	0.008
					$n = 400$					
0.0	1.029	1.020	−0.171	0.118	1.014	2.429	[0.794, 1.265]	0.471	0.956	0.005
1.0	1.053	1.045	−0.157	0.121	1.015	2.474	[0.814, 1.294]	0.480	0.954	0.005
2.0	1.080	1.070	−0.203	0.124	1.020	2.529	[0.835, 1.326]	0.491	0.954	0.005
3.0	1.107	1.103	−0.083	0.127	1.020	2.583	[0.858, 1.359]	0.501	0.956	0.006
4.0	1.135	1.130	−0.093	0.129	1.017	2.642	[0.882, 1.394]	0.513	0.960	0.006
					$n = 500$					
0.0	1.031	1.032	0.015	0.106	0.952	2.440	[0.824, 1.248]	0.424	0.948	0.004
1.0	1.055	1.052	−0.060	0.109	0.966	2.488	[0.843, 1.275]	0.432	0.936	0.004
2.0	1.078	1.072	−0.130	0.111	0.970	2.537	[0.862, 1.303]	0.441	0.934	0.004
3.0	1.105	1.096	−0.193	0.113	0.974	2.581	[0.882, 1.331]	0.448	0.936	0.004
4.0	1.130	1.124	−0.131	0.115	0.977	2.628	[0.902, 1.359]	0.456	0.934	0.005
					$n = 600$					
0.0	1.034	1.037	0.058	0.098	1.036	2.420	[0.844, 1.229]	0.385	0.946	0.003
1.0	1.056	1.053	−0.060	0.099	1.006	2.459	[0.861, 1.252]	0.391	0.948	0.004
2.0	1.077	1.077	−0.015	0.101	1.005	2.500	[0.879, 1.277]	0.398	0.944	0.004
3.0	1.099	1.097	−0.067	0.103	0.998	2.540	[0.898, 1.302]	0.404	0.942	0.004
4.0	1.121	1.120	−0.035	0.105	1.007	2.583	[0.917, 1.328]	0.411	0.948	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Local misspecification v_{ij} is drawn from a $N(0, \sigma)$. [3] Largest degree of misspecification $M = 4$. [4] Sieves dimension $k_n = 3$.

B Additional Tables

Table 10: $\hat{\beta}_n$ Robust Estimator with with $v_{ij} \sim \text{Logistic}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.465	1.368	0.959	0.923	17.060	[−0.517, 3.291]	3.808	0.924	0.010
1.0	1.386	1.358	−0.498	0.919	0.889	16.441	[−0.479, 3.188]	3.667	0.916	0.011
2.0	1.386	1.407	0.356	0.881	0.883	15.724	[−0.394, 3.117]	3.511	0.916	0.012
3.0	1.386	1.405	0.326	0.842	0.893	15.045	[−0.316, 3.046]	3.362	0.918	0.013
4.0	1.386	1.410	0.407	0.816	0.876	14.389	[−0.241, 2.973]	3.214	0.910	0.014
					$n = 400$					
0.0	1.386	1.332	−1.092	0.843	0.925	17.343	[−0.387, 2.977]	3.364	0.932	0.007
1.0	1.386	1.306	−1.608	0.819	0.914	16.675	[−0.347, 2.889]	3.237	0.928	0.008
2.0	1.386	1.310	−1.530	0.787	0.912	16.101	[−0.285, 2.841]	3.126	0.936	0.009
3.0	1.386	1.310	−1.528	0.763	0.901	15.561	[−0.237, 2.784]	3.021	0.926	0.009
4.0	1.386	1.325	−1.222	0.733	0.921	14.993	[−0.187, 2.724]	2.911	0.922	0.010
					$n = 500$					
0.0	1.386	1.319	−1.509	0.768	0.987	17.385	[−0.197, 2.827]	3.024	0.944	0.006
1.0	1.386	1.281	−2.348	0.741	0.990	16.869	[−0.188, 2.745]	2.933	0.940	0.006
2.0	1.386	1.283	−2.302	0.717	0.984	16.328	[−0.150, 2.689]	2.838	0.940	0.007
3.0	1.386	1.291	−2.127	0.695	0.981	15.766	[−0.088, 2.652]	2.740	0.938	0.007
4.0	1.386	1.323	−1.422	0.670	0.975	15.214	[−0.029, 2.616]	2.645	0.938	0.008
					$n = 600$					
0.0	1.386	1.386	−0.001	0.695	0.988	17.243	[−0.014, 2.726]	2.740	0.944	0.005
1.0	1.386	1.374	−0.312	0.671	0.996	16.694	[0.025, 2.680]	2.654	0.956	0.005
2.0	1.386	1.392	0.142	0.656	0.999	16.250	[0.070, 2.654]	2.584	0.950	0.006
3.0	1.386	1.365	−0.523	0.637	1.010	15.866	[0.094, 2.616]	2.523	0.962	0.006
4.0	1.386	1.378	−0.210	0.618	1.022	15.400	[0.132, 2.581]	2.449	0.950	0.006

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 11: $\hat{\beta}_n$ Logistic Estimator with $v_{ij} \sim$ Logistic$(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.481	1.635	0.983	0.935	17.465	[−0.547, 3.353]	3.900	0.926	0.010
1.0	1.386	1.374	−0.211	0.944	0.901	16.868	[−0.510, 3.253]	3.763	0.920	0.011
2.0	1.386	1.425	0.674	0.905	0.894	16.177	[−0.426, 3.186]	3.613	0.920	0.012
3.0	1.386	1.425	0.676	0.869	0.906	15.527	[−0.350, 3.120]	3.470	0.924	0.013
4.0	1.386	1.433	0.803	0.843	0.890	14.903	[−0.276, 3.054]	3.330	0.914	0.014
					$n = 400$					
0.0	1.386	1.343	−0.871	0.857	0.932	17.645	[−0.406, 3.017]	3.423	0.938	0.007
1.0	1.386	1.317	−1.378	0.833	0.922	16.991	[−0.367, 2.932]	3.298	0.930	0.008
2.0	1.386	1.322	−1.279	0.803	0.921	16.433	[−0.305, 2.886]	3.190	0.938	0.009
3.0	1.386	1.323	−1.273	0.781	0.911	15.912	[−0.258, 2.832]	3.090	0.930	0.009
4.0	1.386	1.340	−0.927	0.750	0.931	15.364	[−0.208, 2.775]	2.983	0.922	0.010
					$n = 500$					
0.0	1.386	1.327	−1.318	0.779	0.994	17.626	[−0.209, 2.857]	3.066	0.946	0.006
1.0	1.386	1.291	−2.140	0.753	0.997	17.119	[−0.200, 2.776]	2.977	0.942	0.006
2.0	1.386	1.293	−2.094	0.729	0.992	16.589	[−0.163, 2.721]	2.884	0.942	0.007
3.0	1.386	1.300	−1.923	0.707	0.990	16.039	[−0.101, 2.686]	2.788	0.940	0.007
4.0	1.386	1.334	−1.176	0.683	0.985	15.501	[−0.043, 2.652]	2.695	0.940	0.008
					$n = 600$					
0.0	1.386	1.393	0.166	0.704	0.995	17.442	[−0.022, 2.750]	2.772	0.944	0.005
1.0	1.386	1.381	−0.119	0.679	1.002	16.901	[0.017, 2.704]	2.687	0.960	0.005
2.0	1.386	1.400	0.346	0.664	1.006	16.467	[0.061, 2.680]	2.619	0.950	0.006
3.0	1.386	1.374	−0.301	0.645	1.016	16.093	[0.085, 2.644]	2.559	0.962	0.006
4.0	1.386	1.388	0.034	0.628	1.031	15.638	[0.123, 2.610]	2.487	0.954	0.006

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 12: $\hat{\beta}_n$ Poisson Estimator with $v_{ij} \sim \text{Logistic}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.465	1.366	0.972	0.935	17.256	[−0.539, 3.313]	3.852	0.928	0.010
1.0	1.386	1.357	−0.502	0.931	0.900	16.647	[−0.502, 3.211]	3.713	0.918	0.011
2.0	1.386	1.407	0.351	0.893	0.896	15.942	[−0.419, 3.141]	3.559	0.920	0.012
3.0	1.386	1.405	0.321	0.856	0.907	15.277	[−0.343, 3.071]	3.414	0.926	0.013
4.0	1.386	1.409	0.400	0.829	0.890	14.636	[−0.269, 3.000]	3.269	0.916	0.014
					$n = 400$					
0.0	1.386	1.332	−1.093	0.851	0.933	17.490	[−0.402, 2.991]	3.393	0.936	0.007
1.0	1.386	1.306	−1.610	0.826	0.922	16.829	[−0.362, 2.904]	3.267	0.930	0.008
2.0	1.386	1.310	−1.532	0.795	0.921	16.262	[−0.301, 2.856]	3.157	0.938	0.009
3.0	1.386	1.310	−1.530	0.771	0.911	15.731	[−0.254, 2.801]	3.054	0.930	0.009
4.0	1.386	1.325	−1.226	0.740	0.931	15.173	[−0.204, 2.741]	2.946	0.924	0.010
					$n = 500$					
0.0	1.386	1.319	−1.511	0.774	0.994	17.503	[−0.207, 2.837]	3.044	0.948	0.006
1.0	1.386	1.281	−2.350	0.747	0.998	16.991	[−0.198, 2.756]	2.954	0.942	0.006
2.0	1.386	1.283	−2.303	0.723	0.991	16.455	[−0.161, 2.700]	2.861	0.942	0.007
3.0	1.386	1.291	−2.129	0.700	0.989	15.899	[−0.100, 2.664]	2.763	0.942	0.007
4.0	1.386	1.323	−1.424	0.676	0.985	15.354	[−0.042, 2.628]	2.669	0.938	0.008
					$n = 600$					
0.0	1.386	1.386	−0.002	0.699	0.994	17.340	[−0.021, 2.734]	2.755	0.944	0.005
1.0	1.386	1.374	−0.313	0.675	1.001	16.795	[0.017, 2.688]	2.670	0.960	0.005
2.0	1.386	1.392	0.141	0.660	1.006	16.356	[0.061, 2.662]	2.601	0.950	0.006
3.0	1.386	1.365	−0.525	0.641	1.017	15.976	[0.085, 2.625]	2.540	0.962	0.006
4.0	1.386	1.378	−0.212	0.623	1.030	15.516	[0.123, 2.590]	2.468	0.952	0.006

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 13: $\hat{\beta}_n$ Robust Estimator with $v_{ij} \sim$ Laplace$(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.465	1.368	0.959	0.923	17.060	[−0.517, 3.291]	3.808	0.924	0.010
1.0	1.386	1.449	1.079	0.932	0.907	16.656	[−0.494, 3.223]	3.717	0.918	0.010
2.0	1.386	1.396	0.176	0.904	0.879	16.254	[−0.446, 3.180]	3.626	0.918	0.011
3.0	1.386	1.361	−0.438	0.878	0.881	15.807	[−0.413, 3.114]	3.527	0.916	0.012
4.0	1.386	1.394	0.135	0.859	0.883	15.334	[−0.359, 3.066]	3.426	0.922	0.012
					$n = 400$					
0.0	1.386	1.332	−1.092	0.843	0.925	17.343	[−0.387, 2.977]	3.364	0.932	0.007
1.0	1.386	1.358	−0.562	0.825	0.915	16.872	[−0.333, 2.941]	3.274	0.932	0.008
2.0	1.386	1.321	−1.301	0.806	0.906	16.493	[−0.317, 2.885]	3.202	0.930	0.008
3.0	1.386	1.336	−1.008	0.789	0.888	16.125	[−0.283, 2.848]	3.130	0.918	0.008
4.0	1.386	1.314	−1.438	0.773	0.899	15.784	[−0.250, 2.815]	3.064	0.924	0.009
					$n = 500$					
0.0	1.386	1.319	−1.509	0.768	0.987	17.385	[−0.197, 2.827]	3.024	0.944	0.006
1.0	1.386	1.305	−1.815	0.755	0.991	17.106	[−0.201, 2.773]	2.974	0.936	0.006
2.0	1.386	1.295	−2.034	0.736	0.987	16.705	[−0.165, 2.740]	2.905	0.940	0.006
3.0	1.386	1.291	−2.133	0.719	0.981	16.384	[−0.151, 2.698]	2.849	0.942	0.007
4.0	1.386	1.250	−3.038	0.701	0.982	16.003	[−0.114, 2.668]	2.782	0.944	0.007
					$n = 600$					
0.0	1.386	1.386	−0.001	0.695	0.988	17.243	[−0.014, 2.726]	2.740	0.944	0.005
1.0	1.386	1.365	−0.516	0.677	0.997	16.877	[0.025, 2.708]	2.683	0.948	0.005
2.0	1.386	1.386	−0.010	0.665	0.988	16.600	[0.033, 2.672]	2.639	0.942	0.005
3.0	1.386	1.356	−0.743	0.657	1.010	16.339	[0.048, 2.646]	2.598	0.954	0.005
4.0	1.386	1.369	−0.433	0.645	1.013	16.026	[0.075, 2.624]	2.548	0.956	0.006

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 14: $\hat{\beta}_n$ Logistic Estimator with $v_{ij} \sim \text{Laplace}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.481	1.635	0.983	0.935	17.465	[−0.547, 3.353]	3.900	0.926	0.010
1.0	1.386	1.465	1.356	0.956	0.919	17.074	[−0.525, 3.287]	3.811	0.920	0.010
2.0	1.386	1.414	0.486	0.929	0.892	16.689	[−0.477, 3.247]	3.724	0.918	0.011
3.0	1.386	1.380	−0.115	0.905	0.895	16.258	[−0.445, 3.183]	3.629	0.926	0.012
4.0	1.386	1.414	0.481	0.885	0.896	15.803	[−0.392, 3.139]	3.531	0.928	0.012
					$n = 400$					
0.0	1.386	1.343	−0.871	0.857	0.932	17.645	[−0.406, 3.017]	3.423	0.938	0.007
1.0	1.386	1.370	−0.328	0.839	0.922	17.181	[−0.352, 2.982]	3.335	0.932	0.008
2.0	1.386	1.332	−1.078	0.822	0.914	16.812	[−0.336, 2.928]	3.265	0.932	0.008
3.0	1.386	1.349	−0.743	0.807	0.900	16.455	[−0.302, 2.893]	3.195	0.922	0.008
4.0	1.386	1.328	−1.176	0.790	0.908	16.127	[−0.270, 2.862]	3.132	0.930	0.009
					$n = 500$					
0.0	1.386	1.327	−1.318	0.779	0.994	17.626	[−0.209, 2.857]	3.066	0.946	0.006
1.0	1.386	1.315	−1.603	0.766	0.998	17.354	[−0.214, 2.803]	3.017	0.940	0.006
2.0	1.386	1.304	−1.831	0.747	0.993	16.958	[−0.178, 2.771]	2.949	0.946	0.006
3.0	1.386	1.300	−1.919	0.730	0.988	16.644	[−0.164, 2.731]	2.894	0.942	0.007
4.0	1.386	1.260	−2.827	0.713	0.990	16.271	[−0.128, 2.701]	2.829	0.946	0.007
					$n = 600$					
0.0	1.386	1.393	0.166	0.704	0.995	17.442	[−0.022, 2.750]	2.772	0.944	0.005
1.0	1.386	1.373	−0.333	0.684	1.003	17.080	[0.016, 2.732]	2.715	0.948	0.005
2.0	1.386	1.394	0.192	0.673	0.995	16.810	[0.024, 2.697]	2.672	0.946	0.005
3.0	1.386	1.364	−0.536	0.666	1.017	16.555	[0.039, 2.672]	2.632	0.954	0.005
4.0	1.386	1.377	−0.217	0.654	1.020	16.248	[0.066, 2.650]	2.584	0.958	0.006

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 15: $\hat{\beta}_n$ Poisson Estimator with $v_{ij} \sim \text{Laplace}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.465	1.366	0.972	0.935	17.256	[−0.539, 3.313]	3.852	0.928	0.010
1.0	1.386	1.448	1.076	0.942	0.918	16.858	[−0.517, 3.245]	3.762	0.924	0.010
2.0	1.386	1.396	0.172	0.917	0.892	16.464	[−0.470, 3.204]	3.673	0.918	0.011
3.0	1.386	1.361	−0.443	0.892	0.894	16.025	[−0.438, 3.138]	3.576	0.922	0.012
4.0	1.386	1.394	0.130	0.871	0.896	15.560	[−0.385, 3.091]	3.476	0.932	0.012
					$n = 400$					
0.0	1.386	1.332	−1.093	0.851	0.933	17.490	[−0.402, 2.991]	3.393	0.936	0.007
1.0	1.386	1.358	−0.564	0.832	0.922	17.022	[−0.348, 2.955]	3.303	0.934	0.008
2.0	1.386	1.321	−1.303	0.814	0.914	16.648	[−0.332, 2.900]	3.232	0.930	0.008
3.0	1.386	1.336	−1.010	0.798	0.898	16.285	[−0.298, 2.863]	3.162	0.922	0.008
4.0	1.386	1.314	−1.440	0.781	0.908	15.950	[−0.266, 2.831]	3.097	0.930	0.009
					$n = 500$					
0.0	1.386	1.319	−1.511	0.774	0.994	17.503	[−0.207, 2.837]	3.044	0.948	0.006
1.0	1.386	1.305	−1.816	0.760	0.998	17.227	[−0.212, 2.783]	2.995	0.940	0.006
2.0	1.386	1.295	−2.036	0.741	0.993	16.828	[−0.176, 2.751]	2.926	0.942	0.006
3.0	1.386	1.291	−2.134	0.725	0.988	16.511	[−0.162, 2.709]	2.871	0.942	0.007
4.0	1.386	1.250	−3.040	0.707	0.990	16.133	[−0.126, 2.679]	2.805	0.946	0.007
					$n = 600$					
0.0	1.386	1.386	−0.002	0.699	0.994	17.340	[−0.021, 2.734]	2.755	0.944	0.005
1.0	1.386	1.365	−0.517	0.680	1.003	16.976	[0.017, 2.716]	2.699	0.948	0.005
2.0	1.386	1.386	−0.012	0.669	0.995	16.702	[0.025, 2.680]	2.655	0.946	0.005
3.0	1.386	1.356	−0.745	0.662	1.017	16.444	[0.040, 2.654]	2.615	0.954	0.005
4.0	1.386	1.369	−0.435	0.649	1.020	16.134	[0.067, 2.632]	2.566	0.958	0.006

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 16: $\hat{\beta}_n$ Robust Estimator with $v_{ij} \sim \mathrm{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.465	1.368	0.959	0.923	17.060	[−0.517, 3.291]	3.808	0.924	0.010
1.0	1.386	1.446	1.037	0.949	0.913	16.826	[−0.510, 3.245]	3.755	0.926	0.010
2.0	1.386	1.417	0.525	0.938	0.911	16.673	[−0.505, 3.215]	3.721	0.922	0.010
3.0	1.386	1.384	−0.034	0.931	0.910	16.545	[−0.487, 3.203]	3.690	0.916	0.010
4.0	1.386	1.378	−0.140	0.912	0.903	16.294	[−0.465, 3.169]	3.634	0.914	0.011
					$n = 400$					
0.0	1.386	1.332	−1.092	0.843	0.925	17.343	[−0.387, 2.977]	3.364	0.932	0.007
1.0	1.386	1.337	−0.984	0.831	0.918	17.105	[−0.367, 2.952]	3.319	0.928	0.007
2.0	1.386	1.318	−1.360	0.822	0.915	16.901	[−0.360, 2.920]	3.280	0.924	0.008
3.0	1.386	1.317	−1.386	0.817	0.925	16.736	[−0.342, 2.907]	3.249	0.936	0.008
4.0	1.386	1.327	−1.190	0.813	0.924	16.584	[−0.334, 2.885]	3.219	0.940	0.008
					$n = 500$					
0.0	1.386	1.319	−1.509	0.768	0.987	17.385	[−0.197, 2.827]	3.024	0.944	0.006
1.0	1.386	1.267	−2.668	0.763	0.999	17.266	[−0.214, 2.787]	3.002	0.946	0.006
2.0	1.386	1.263	−2.751	0.758	1.005	17.139	[−0.218, 2.763]	2.980	0.948	0.006
3.0	1.386	1.270	−2.601	0.747	1.000	16.984	[−0.205, 2.747]	2.953	0.946	0.006
4.0	1.386	1.238	−3.319	0.740	0.986	16.803	[−0.195, 2.725]	2.920	0.936	0.006
					$n = 600$					
0.0	1.386	1.386	−0.001	0.695	0.988	17.243	[−0.014, 2.726]	2.740	0.944	0.005
1.0	1.386	1.374	−0.289	0.683	0.993	17.067	[0.001, 2.714]	2.713	0.942	0.005
2.0	1.386	1.377	−0.225	0.681	1.007	16.900	[0.015, 2.701]	2.687	0.958	0.005
3.0	1.386	1.414	0.675	0.675	1.001	16.756	[0.032, 2.696]	2.664	0.956	0.005
4.0	1.386	1.364	−0.557	0.672	1.020	16.624	[0.041, 2.684]	2.643	0.960	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 17: $\hat{\beta}_n$ Logistic Estimator with $v_{ij} \sim \mathrm{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.481	1.635	0.983	0.935	17.465	[−0.547, 3.353]	3.899	0.926	0.010
1.0	1.386	1.462	1.315	0.972	0.924	17.235	[−0.540, 3.307]	3.847	0.926	0.010
2.0	1.386	1.433	0.807	0.963	0.923	17.089	[−0.536, 3.279]	3.815	0.926	0.010
3.0	1.386	1.399	0.215	0.955	0.922	16.970	[−0.518, 3.268]	3.786	0.918	0.010
4.0	1.386	1.395	0.148	0.936	0.915	16.724	[−0.496, 3.235]	3.730	0.916	0.011
					$n = 400$					
0.0	1.386	1.343	−0.871	0.857	0.932	17.645	[−0.406, 3.017]	3.423	0.938	0.007
1.0	1.386	1.348	−0.770	0.846	0.926	17.410	[−0.386, 2.993]	3.378	0.932	0.007
2.0	1.386	1.329	−1.145	0.839	0.925	17.210	[−0.379, 2.962]	3.341	0.928	0.008
3.0	1.386	1.328	−1.157	0.833	0.934	17.050	[−0.361, 2.949]	3.311	0.940	0.008
4.0	1.386	1.338	−0.964	0.828	0.932	16.902	[−0.354, 2.928]	3.281	0.940	0.008
					$n = 500$					
0.0	1.386	1.327	−1.318	0.779	0.994	17.626	[−0.209, 2.857]	3.066	0.946	0.006
1.0	1.386	1.275	−2.484	0.775	1.007	17.509	[−0.227, 2.817]	3.044	0.950	0.006
2.0	1.386	1.272	−2.567	0.768	1.011	17.386	[−0.231, 2.793]	3.024	0.950	0.006
3.0	1.386	1.279	−2.410	0.758	1.007	17.234	[−0.218, 2.778]	2.996	0.946	0.006
4.0	1.386	1.246	−3.133	0.750	0.993	17.055	[−0.208, 2.756]	2.964	0.936	0.006
					$n = 600$					
0.0	1.386	1.393	0.166	0.704	0.995	17.442	[−0.022, 2.750]	2.772	0.944	0.005
1.0	1.386	1.382	−0.111	0.691	0.999	17.268	[−0.007, 2.738]	2.745	0.946	0.005
2.0	1.386	1.384	−0.046	0.689	1.013	17.104	[0.006, 2.725]	2.719	0.960	0.005
3.0	1.386	1.422	0.866	0.683	1.006	16.962	[0.024, 2.720]	2.697	0.956	0.005
4.0	1.386	1.371	−0.365	0.680	1.026	16.833	[0.032, 2.708]	2.676	0.960	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 18: $\hat{\beta}_n$ Poisson Estimator with $v_{ij} \sim \mathrm{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.386	1.465	1.366	0.972	0.935	17.256	[−0.539, 3.313]	3.852	0.928	0.010
1.0	1.386	1.446	1.034	0.960	0.924	17.023	[−0.532, 3.267]	3.799	0.926	0.010
2.0	1.386	1.416	0.521	0.950	0.922	16.874	[−0.528, 3.238]	3.766	0.924	0.010
3.0	1.386	1.384	−0.036	0.942	0.921	16.751	[−0.510, 3.226]	3.736	0.920	0.010
4.0	1.386	1.378	−0.143	0.923	0.915	16.502	[−0.488, 3.192]	3.680	0.916	0.011
					$n = 400$					
0.0	1.386	1.332	−1.093	0.851	0.933	17.490	[−0.402, 2.991]	3.393	0.936	0.007
1.0	1.386	1.337	−0.986	0.838	0.926	17.253	[−0.381, 2.966]	3.347	0.928	0.007
2.0	1.386	1.318	−1.362	0.829	0.924	17.051	[−0.375, 2.935]	3.309	0.926	0.008
3.0	1.386	1.317	−1.388	0.825	0.934	16.889	[−0.357, 2.922]	3.279	0.940	0.008
4.0	1.386	1.327	−1.192	0.820	0.933	16.739	[−0.349, 2.900]	3.249	0.942	0.008
					$n = 500$					
0.0	1.386	1.319	−1.511	0.774	0.994	17.503	[−0.207, 2.837]	3.044	0.948	0.006
1.0	1.386	1.267	−2.669	0.768	1.006	17.385	[−0.225, 2.798]	3.022	0.950	0.006
2.0	1.386	1.263	−2.752	0.763	1.011	17.260	[−0.228, 2.773]	3.001	0.952	0.006
3.0	1.386	1.270	−2.602	0.753	1.007	17.106	[−0.216, 2.758]	2.974	0.946	0.006
4.0	1.386	1.238	−3.321	0.745	0.993	16.926	[−0.206, 2.736]	2.942	0.936	0.006
					$n = 600$					
0.0	1.386	1.386	−0.002	0.699	0.994	17.340	[−0.021, 2.734]	2.755	0.944	0.005
1.0	1.386	1.374	−0.290	0.687	0.999	17.165	[−0.007, 2.722]	2.729	0.948	0.005
2.0	1.386	1.377	−0.226	0.685	1.013	16.999	[0.007, 2.709]	2.702	0.960	0.005
3.0	1.386	1.414	0.674	0.680	1.007	16.856	[0.024, 2.704]	2.680	0.956	0.005
4.0	1.386	1.364	−0.558	0.676	1.026	16.726	[0.033, 2.692]	2.659	0.960	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 2$.

Table 19: $\hat{\Psi}_n$ Robust Estimator with with $v_{ij} \sim \text{Logistic}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.007	-0.444	0.130	0.884	2.353	[0.754, 1.282]	0.528	0.910	0.007
1.0	1.139	1.117	-0.377	0.142	0.901	2.552	[0.837, 1.410]	0.573	0.910	0.008
2.0	1.257	1.234	-0.386	0.154	0.890	2.773	[0.932, 1.554]	0.622	0.920	0.008
3.0	1.393	1.368	-0.419	0.168	0.899	3.010	[1.040, 1.715]	0.676	0.910	0.009
4.0	1.544	1.515	-0.509	0.182	0.898	3.253	[1.160, 1.889]	0.730	0.916	0.010
					$n = 400$					
0.0	1.036	1.009	-0.536	0.115	0.958	2.364	[0.787, 1.248]	0.461	0.930	0.005
1.0	1.125	1.105	-0.411	0.124	0.951	2.539	[0.861, 1.355]	0.494	0.928	0.006
2.0	1.226	1.205	-0.414	0.133	0.938	2.730	[0.947, 1.478]	0.531	0.932	0.006
3.0	1.341	1.323	-0.359	0.144	0.941	2.935	[1.040, 1.611]	0.571	0.934	0.007
4.0	1.467	1.446	-0.431	0.154	0.949	3.154	[1.143, 1.756]	0.614	0.926	0.007
					$n = 500$					
0.0	1.033	1.005	-0.622	0.103	0.923	2.365	[0.810, 1.222]	0.412	0.926	0.004
1.0	1.112	1.092	-0.436	0.111	0.947	2.514	[0.879, 1.317]	0.438	0.932	0.004
2.0	1.202	1.175	-0.605	0.118	0.951	2.680	[0.952, 1.418]	0.467	0.924	0.005
3.0	1.302	1.277	-0.561	0.126	0.940	2.859	[1.034, 1.532]	0.498	0.916	0.005
4.0	1.412	1.389	-0.503	0.134	0.944	3.059	[1.125, 1.658]	0.533	0.920	0.006
					$n = 600$					
0.0	1.027	1.015	-0.283	0.095	1.038	2.380	[0.832, 1.211]	0.379	0.958	0.003
1.0	1.098	1.090	-0.198	0.100	1.039	2.525	[0.894, 1.296]	0.402	0.958	0.004
2.0	1.176	1.167	-0.223	0.106	1.019	2.675	[0.960, 1.386]	0.426	0.954	0.004
3.0	1.264	1.258	-0.150	0.113	1.026	2.846	[1.035, 1.488]	0.453	0.954	0.004
4.0	1.362	1.354	-0.198	0.121	1.033	3.031	[1.118, 1.600]	0.482	0.956	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 20: $\hat{\Psi}_n$ Logistic Estimator with $v_{ij} \sim$ Logistic$(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.050	0.295	0.140	0.934	2.549	[0.761, 1.362]	0.602	0.958	0.007
1.0	1.139	1.163	0.413	0.153	0.950	2.785	[0.844, 1.498]	0.654	0.956	0.008
2.0	1.257	1.286	0.503	0.167	0.945	3.037	[0.942, 1.652]	0.710	0.960	0.008
3.0	1.393	1.438	0.781	0.183	0.948	3.323	[1.053, 1.827]	0.774	0.958	0.009
4.0	1.544	1.590	0.784	0.200	0.946	3.631	[1.177, 2.019]	0.842	0.958	0.010
					$n = 400$					
0.0	1.036	1.044	0.160	0.122	1.001	2.503	[0.794, 1.306]	0.513	0.958	0.005
1.0	1.125	1.139	0.281	0.131	0.993	2.696	[0.871, 1.421]	0.550	0.960	0.006
2.0	1.226	1.249	0.450	0.142	0.988	2.908	[0.959, 1.550]	0.591	0.958	0.006
3.0	1.341	1.369	0.552	0.153	0.983	3.139	[1.054, 1.690]	0.636	0.952	0.007
4.0	1.467	1.498	0.617	0.166	1.001	3.393	[1.159, 1.844]	0.685	0.960	0.007
					$n = 500$					
0.0	1.033	1.030	-0.082	0.108	0.974	2.491	[0.815, 1.269]	0.453	0.960	0.004
1.0	1.112	1.126	0.323	0.116	0.985	2.663	[0.885, 1.368]	0.483	0.956	0.004
2.0	1.202	1.216	0.324	0.124	0.990	2.843	[0.961, 1.475]	0.514	0.956	0.005
3.0	1.302	1.313	0.242	0.132	0.983	3.047	[1.044, 1.593]	0.549	0.950	0.005
4.0	1.412	1.426	0.318	0.142	0.989	3.271	[1.138, 1.726]	0.588	0.954	0.006
					$n = 600$					
0.0	1.027	1.041	0.355	0.099	1.065	2.483	[0.838, 1.251]	0.413	0.976	0.003
1.0	1.098	1.116	0.425	0.105	1.056	2.641	[0.902, 1.339]	0.438	0.980	0.004
2.0	1.176	1.194	0.448	0.111	1.042	2.804	[0.970, 1.433]	0.464	0.982	0.004
3.0	1.264	1.289	0.592	0.119	1.048	2.992	[1.047, 1.541]	0.493	0.972	0.004
4.0	1.362	1.390	0.684	0.127	1.047	3.193	[1.131, 1.656]	0.525	0.978	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 21: $\hat{\Psi}_n$ Poisson Estimator with $v_{ij} \sim \text{Logistic}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.033	0.005	0.135	0.945	2.425	[0.771, 1.312]	0.541	0.938	0.007
1.0	1.139	1.140	0.014	0.147	0.963	2.632	[0.853, 1.440]	0.588	0.944	0.008
2.0	1.257	1.260	0.058	0.159	0.957	2.853	[0.949, 1.586]	0.637	0.944	0.008
3.0	1.393	1.403	0.186	0.174	0.962	3.103	[1.057, 1.750]	0.694	0.932	0.009
4.0	1.544	1.550	0.098	0.190	0.964	3.367	[1.178, 1.930]	0.752	0.942	0.010
					$n = 400$					
0.0	1.036	1.030	−0.115	0.118	1.006	2.426	[0.800, 1.271]	0.471	0.956	0.005
1.0	1.125	1.125	−0.008	0.127	1.001	2.603	[0.876, 1.381]	0.505	0.950	0.006
2.0	1.226	1.230	0.077	0.137	0.996	2.797	[0.962, 1.505]	0.543	0.948	0.006
3.0	1.341	1.345	0.078	0.147	0.992	3.007	[1.055, 1.639]	0.584	0.950	0.007
4.0	1.467	1.470	0.055	0.158	1.004	3.235	[1.158, 1.787]	0.629	0.946	0.007
					$n = 500$					
0.0	1.033	1.020	−0.298	0.106	0.981	2.422	[0.820, 1.241]	0.421	0.944	0.004
1.0	1.112	1.112	0.004	0.113	0.992	2.581	[0.889, 1.338]	0.449	0.940	0.004
2.0	1.202	1.200	−0.041	0.121	0.999	2.749	[0.963, 1.441]	0.478	0.942	0.005
3.0	1.302	1.296	−0.133	0.129	0.993	2.933	[1.045, 1.555]	0.510	0.944	0.005
4.0	1.412	1.406	−0.125	0.138	0.998	3.140	[1.138, 1.683]	0.546	0.934	0.006
					$n = 600$					
0.0	1.027	1.032	0.121	0.098	1.071	2.431	[0.842, 1.228]	0.386	0.968	0.003
1.0	1.098	1.103	0.122	0.103	1.064	2.581	[0.904, 1.314]	0.410	0.968	0.004
2.0	1.176	1.183	0.179	0.109	1.050	2.734	[0.971, 1.406]	0.435	0.966	0.004
3.0	1.264	1.277	0.301	0.116	1.056	2.910	[1.048, 1.510]	0.462	0.956	0.004
4.0	1.362	1.373	0.284	0.124	1.054	3.097	[1.130, 1.622]	0.492	0.956	0.005

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 22: $\hat{\Psi}_n$ Robust Estimator with $v_{ij} \sim \text{Laplace}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.007	−0.444	0.130	0.884	2.353	[0.755, 1.282]	0.527	0.910	0.007
1.0	1.096	1.078	−0.316	0.138	0.897	2.480	[0.806, 1.362]	0.556	0.918	0.007
2.0	1.167	1.147	−0.350	0.146	0.904	2.609	[0.862, 1.446]	0.584	0.912	0.008
3.0	1.246	1.225	−0.356	0.153	0.896	2.745	[0.925, 1.540]	0.615	0.912	0.008
4.0	1.332	1.310	−0.390	0.161	0.891	2.893	[0.994, 1.642]	0.648	0.910	0.009
					$n = 400$					
0.0	1.036	1.009	−0.536	0.115	0.958	2.364	[0.787, 1.247]	0.460	0.930	0.005
1.0	1.090	1.067	−0.473	0.121	0.949	2.471	[0.833, 1.313]	0.481	0.926	0.005
2.0	1.150	1.132	−0.363	0.127	0.950	2.588	[0.882, 1.385]	0.503	0.924	0.006
3.0	1.218	1.197	−0.417	0.132	0.940	2.707	[0.938, 1.464]	0.526	0.930	0.006
4.0	1.291	1.275	−0.337	0.139	0.937	2.841	[1.000, 1.552]	0.552	0.928	0.006
					$n = 500$					
0.0	1.033	1.005	−0.622	0.103	0.923	2.365	[0.810, 1.222]	0.412	0.926	0.004
1.0	1.080	1.062	−0.416	0.108	0.935	2.455	[0.852, 1.279]	0.427	0.924	0.004
2.0	1.133	1.113	−0.448	0.111	0.930	2.552	[0.896, 1.340]	0.444	0.920	0.005
3.0	1.192	1.169	−0.516	0.117	0.939	2.661	[0.944, 1.407]	0.463	0.920	0.005
4.0	1.257	1.234	−0.512	0.122	0.946	2.778	[0.997, 1.481]	0.483	0.922	0.005
					$n = 600$					
0.0	1.027	1.015	−0.283	0.095	1.038	2.380	[0.832, 1.211]	0.379	0.956	0.003
1.0	1.070	1.060	−0.249	0.098	1.052	2.472	[0.870, 1.264]	0.393	0.956	0.004
2.0	1.116	1.103	−0.323	0.102	1.035	2.559	[0.909, 1.316]	0.407	0.958	0.004
3.0	1.167	1.159	−0.208	0.106	1.024	2.657	[0.954, 1.377]	0.423	0.956	0.004
4.0	1.224	1.216	−0.194	0.110	1.028	2.765	[1.002, 1.441]	0.440	0.954	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 23: $\hat{\Psi}_n$ Logistic Estimator with $v_{ij} \sim \text{Laplace}(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.050	0.295	0.140	0.934	2.549	[0.761, 1.362]	0.602	0.958	0.007
1.0	1.096	1.122	0.443	0.147	0.938	2.699	[0.812, 1.446]	0.634	0.960	0.007
2.0	1.167	1.193	0.450	0.156	0.946	2.851	[0.868, 1.537]	0.668	0.956	0.008
3.0	1.246	1.276	0.527	0.165	0.949	3.004	[0.934, 1.636]	0.703	0.956	0.008
4.0	1.332	1.367	0.603	0.176	0.955	3.178	[1.005, 1.747]	0.742	0.962	0.009
					$n = 400$					
0.0	1.036	1.044	0.160	0.122	1.001	2.503	[0.794, 1.306]	0.513	0.958	0.005
1.0	1.090	1.100	0.197	0.127	0.984	2.618	[0.841, 1.376]	0.535	0.962	0.005
2.0	1.150	1.167	0.351	0.134	0.990	2.746	[0.892, 1.452]	0.560	0.962	0.006
3.0	1.218	1.237	0.385	0.140	0.981	2.882	[0.949, 1.535]	0.586	0.956	0.006
4.0	1.291	1.317	0.516	0.148	0.978	3.031	[1.012, 1.627]	0.615	0.960	0.006
					$n = 500$					
0.0	1.033	1.030	−0.082	0.108	0.974	2.491	[0.815, 1.269]	0.453	0.960	0.004
1.0	1.080	1.091	0.227	0.113	0.978	2.597	[0.857, 1.329]	0.471	0.958	0.004
2.0	1.133	1.142	0.197	0.118	0.981	2.704	[0.903, 1.393]	0.490	0.952	0.005
3.0	1.192	1.202	0.213	0.123	0.981	2.825	[0.952, 1.463]	0.511	0.952	0.005
4.0	1.257	1.266	0.210	0.129	0.993	2.953	[1.007, 1.540]	0.533	0.954	0.005
					$n = 600$					
0.0	1.027	1.041	0.355	0.099	1.065	2.483	[0.838, 1.251]	0.413	0.976	0.003
1.0	1.070	1.087	0.411	0.103	1.069	2.583	[0.878, 1.306]	0.429	0.978	0.004
2.0	1.116	1.131	0.352	0.107	1.058	2.680	[0.917, 1.361]	0.444	0.980	0.004
3.0	1.167	1.192	0.616	0.111	1.048	2.786	[0.963, 1.424]	0.461	0.980	0.004
4.0	1.224	1.250	0.647	0.115	1.050	2.902	[1.012, 1.491]	0.479	0.976	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 24: $\hat{\Psi}_n$ Poisson Estimator with $v_{ij} \sim$ Laplace$(0, \sigma)$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.033	0.005	0.135	0.945	2.425	[0.771, 1.312]	0.541	0.938	0.007
1.0	1.096	1.100	0.061	0.143	0.956	2.555	[0.821, 1.391]	0.570	0.942	0.007
2.0	1.167	1.173	0.103	0.149	0.962	2.690	[0.876, 1.477]	0.601	0.936	0.008
3.0	1.246	1.253	0.127	0.157	0.959	2.825	[0.940, 1.571]	0.631	0.934	0.008
4.0	1.332	1.337	0.075	0.166	0.961	2.980	[1.009, 1.676]	0.666	0.940	0.009
					$n = 400$					
0.0	1.036	1.030	−0.115	0.118	1.006	2.426	[0.800, 1.271]	0.471	0.956	0.005
1.0	1.090	1.085	−0.104	0.123	0.995	2.531	[0.846, 1.338]	0.491	0.948	0.005
2.0	1.150	1.150	0.004	0.130	0.995	2.649	[0.897, 1.411]	0.514	0.944	0.006
3.0	1.218	1.215	−0.050	0.136	0.992	2.773	[0.953, 1.491]	0.539	0.950	0.006
4.0	1.291	1.295	0.071	0.143	0.987	2.910	[1.014, 1.579]	0.565	0.952	0.006
					$n = 500$					
0.0	1.033	1.020	−0.298	0.106	0.981	2.422	[0.820, 1.241]	0.421	0.944	0.004
1.0	1.080	1.076	−0.100	0.110	0.988	2.518	[0.862, 1.299]	0.438	0.944	0.004
2.0	1.133	1.128	−0.111	0.115	0.990	2.619	[0.906, 1.361]	0.455	0.938	0.005
3.0	1.192	1.188	−0.091	0.119	0.987	2.730	[0.955, 1.429]	0.474	0.946	0.005
4.0	1.257	1.248	−0.195	0.125	1.000	2.849	[1.009, 1.504]	0.495	0.946	0.005
					$n = 600$					
0.0	1.027	1.032	0.121	0.098	1.071	2.431	[0.842, 1.228]	0.386	0.968	0.003
1.0	1.070	1.077	0.165	0.101	1.076	2.525	[0.880, 1.282]	0.401	0.964	0.004
2.0	1.116	1.117	0.012	0.104	1.066	2.617	[0.920, 1.336]	0.416	0.972	0.004
3.0	1.167	1.180	0.313	0.108	1.055	2.717	[0.965, 1.397]	0.432	0.966	0.004
4.0	1.224	1.237	0.320	0.113	1.056	2.826	[1.013, 1.462]	0.449	0.966	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 25: $\hat{\Psi}_n$ Robust Estimator with $v_{ij} \sim \mathrm{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.007	−0.444	0.130	0.884	2.353	[0.756, 1.281]	0.525	0.910	0.007
1.0	1.062	1.040	−0.395	0.133	0.889	2.410	[0.779, 1.317]	0.538	0.912	0.007
2.0	1.094	1.071	−0.396	0.138	0.905	2.472	[0.804, 1.356]	0.552	0.908	0.007
3.0	1.124	1.101	−0.406	0.141	0.907	2.529	[0.828, 1.393]	0.565	0.914	0.008
4.0	1.156	1.132	−0.408	0.144	0.893	2.593	[0.853, 1.432]	0.579	0.908	0.008
					$n = 400$					
0.0	1.036	1.009	−0.536	0.115	0.958	2.364	[0.788, 1.247]	0.459	0.930	0.005
1.0	1.062	1.035	−0.537	0.118	0.958	2.418	[0.810, 1.279]	0.469	0.934	0.005
2.0	1.088	1.065	−0.469	0.121	0.944	2.472	[0.832, 1.312]	0.480	0.922	0.005
3.0	1.114	1.088	−0.514	0.123	0.941	2.522	[0.854, 1.344]	0.490	0.924	0.006
4.0	1.141	1.120	−0.426	0.126	0.956	2.574	[0.878, 1.377]	0.500	0.930	0.006
					$n = 500$					
0.0	1.033	1.005	−0.622	0.103	0.923	2.365	[0.811, 1.222]	0.411	0.926	0.004
1.0	1.055	1.037	−0.388	0.106	0.926	2.407	[0.830, 1.249]	0.419	0.920	0.004
2.0	1.079	1.056	−0.502	0.108	0.933	2.452	[0.850, 1.277]	0.426	0.914	0.004
3.0	1.103	1.084	−0.413	0.109	0.938	2.496	[0.870, 1.304]	0.434	0.916	0.004
4.0	1.127	1.109	−0.382	0.111	0.940	2.543	[0.890, 1.332]	0.442	0.920	0.005
					$n = 600$					
0.0	1.027	1.015	−0.283	0.095	1.038	2.380	[0.832, 1.211]	0.378	0.956	0.003
1.0	1.048	1.036	−0.304	0.097	1.034	2.425	[0.850, 1.235]	0.385	0.950	0.004
2.0	1.070	1.060	−0.247	0.098	1.032	2.467	[0.869, 1.261]	0.392	0.954	0.004
3.0	1.090	1.078	−0.299	0.099	1.018	2.507	[0.887, 1.285]	0.398	0.954	0.004
4.0	1.112	1.103	−0.200	0.101	1.025	2.550	[0.904, 1.310]	0.405	0.958	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 26: $\hat{\Psi}_n$ Logistic Estimator with $v_{ij} \sim \mathrm{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.050	0.295	0.140	0.934	2.549	[0.761, 1.362]	0.602	0.958	0.007
1.0	1.062	1.083	0.352	0.142	0.929	2.617	[0.784, 1.400]	0.616	0.958	0.007
2.0	1.094	1.117	0.406	0.147	0.947	2.686	[0.809, 1.441]	0.632	0.954	0.007
3.0	1.124	1.153	0.498	0.151	0.949	2.755	[0.835, 1.482]	0.647	0.958	0.008
4.0	1.156	1.187	0.531	0.156	0.950	2.831	[0.860, 1.523]	0.664	0.958	0.008
					$n = 400$					
0.0	1.036	1.044	0.160	0.122	1.001	2.503	[0.794, 1.306]	0.513	0.958	0.005
1.0	1.062	1.071	0.181	0.124	0.997	2.561	[0.817, 1.340]	0.524	0.968	0.005
2.0	1.088	1.102	0.280	0.128	0.983	2.623	[0.840, 1.376]	0.536	0.960	0.005
3.0	1.114	1.131	0.327	0.130	0.987	2.678	[0.863, 1.410]	0.546	0.960	0.006
4.0	1.141	1.164	0.460	0.132	0.987	2.733	[0.886, 1.443]	0.557	0.956	0.006
					$n = 500$					
0.0	1.033	1.030	−0.082	0.108	0.974	2.491	[0.815, 1.269]	0.453	0.960	0.004
1.0	1.055	1.065	0.224	0.111	0.968	2.545	[0.836, 1.299]	0.463	0.952	0.004
2.0	1.079	1.089	0.218	0.113	0.972	2.596	[0.857, 1.328]	0.471	0.954	0.004
3.0	1.103	1.119	0.369	0.115	0.986	2.644	[0.877, 1.357]	0.480	0.958	0.004
4.0	1.127	1.141	0.325	0.117	0.985	2.692	[0.898, 1.386]	0.488	0.958	0.005
					$n = 600$					
0.0	1.027	1.041	0.355	0.099	1.065	2.483	[0.838, 1.251]	0.413	0.976	0.003
1.0	1.048	1.062	0.327	0.101	1.055	2.531	[0.857, 1.277]	0.420	0.974	0.004
2.0	1.070	1.083	0.329	0.102	1.050	2.578	[0.876, 1.304]	0.428	0.974	0.004
3.0	1.090	1.106	0.386	0.104	1.044	2.620	[0.894, 1.329]	0.434	0.976	0.004
4.0	1.112	1.130	0.446	0.106	1.052	2.669	[0.912, 1.355]	0.442	0.978	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

Table 27: $\hat{\Psi}_n$ Poisson Estimator with $v_{ij} \sim \mathrm{U}\left[-\sqrt{3}\sigma, \sqrt{3}\sigma\right]$

σ^2	True	coeff.	$\sqrt{n}$-Bias	s.e.	s.e./sd	RMSE	[conf. int.]	width	95% CP	Degree
					$n = 300$					
0.0	1.033	1.033	0.005	0.135	0.945	2.425	[0.771, 1.312]	0.541	0.938	0.007
1.0	1.062	1.060	−0.041	0.137	0.943	2.484	[0.793, 1.348]	0.554	0.940	0.007
2.0	1.094	1.093	−0.012	0.142	0.962	2.544	[0.818, 1.386]	0.568	0.940	0.007
3.0	1.124	1.130	0.104	0.146	0.966	2.606	[0.844, 1.425]	0.582	0.940	0.008
4.0	1.156	1.160	0.071	0.149	0.960	2.671	[0.868, 1.464]	0.596	0.944	0.008
					$n = 400$					
0.0	1.036	1.030	−0.115	0.118	1.006	2.426	[0.800, 1.271]	0.471	0.956	0.005
1.0	1.062	1.055	−0.146	0.121	1.008	2.479	[0.823, 1.304]	0.481	0.950	0.005
2.0	1.088	1.085	−0.060	0.124	0.991	2.536	[0.846, 1.338]	0.492	0.940	0.005
3.0	1.114	1.115	0.017	0.127	0.994	2.587	[0.868, 1.370]	0.502	0.946	0.006
4.0	1.141	1.145	0.079	0.129	0.996	2.637	[0.891, 1.403]	0.512	0.946	0.006
					$n = 500$					
0.0	1.033	1.020	−0.298	0.106	0.981	2.422	[0.820, 1.241]	0.421	0.944	0.004
1.0	1.055	1.052	−0.057	0.108	0.975	2.471	[0.841, 1.270]	0.429	0.944	0.004
2.0	1.079	1.076	−0.064	0.110	0.980	2.519	[0.861, 1.298]	0.438	0.942	0.004
3.0	1.103	1.108	0.116	0.112	0.994	2.563	[0.881, 1.326]	0.445	0.944	0.004
4.0	1.127	1.128	0.033	0.114	0.994	2.609	[0.901, 1.354]	0.453	0.940	0.005
					$n = 600$					
0.0	1.027	1.032	0.121	0.098	1.071	2.431	[0.842, 1.228]	0.386	0.968	0.003
1.0	1.048	1.053	0.122	0.099	1.061	2.477	[0.860, 1.253]	0.394	0.956	0.004
2.0	1.070	1.073	0.089	0.100	1.059	2.522	[0.879, 1.280]	0.401	0.964	0.004
3.0	1.090	1.097	0.157	0.102	1.051	2.561	[0.897, 1.304]	0.407	0.966	0.004
4.0	1.112	1.118	0.167	0.104	1.059	2.607	[0.915, 1.329]	0.414	0.962	0.004

[1] Number of Monte Carlo simulations is $1,000$. [2] Largest degree of misspecification $M = 4$. [3] Sieves dimension $k_n = 3$.

References

Anderson, K. A. and S. Richards-Shubik (2022): "Collaborative production in science: An empirical analysis of coauthorships in economics," *Review of Economics and Statistics*, 104, 1241–1255.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

——— (2020): "On the informativeness of descriptive statistics for structural estimates," *Econometrica*, 88, 2231–2258.

Armstrong, T. B., P. Kline, and L. Sun (2023): "Adapting to Misspecification," *arXiv preprint arXiv:2305.14265*.

Armstrong, T. B. and M. Kolesár (2021): "Sensitivity analysis using approximate moment condition models," *Quantitative Economics*, 12, 77–108.

Armstrong, T. B., M. Weidner, and A. Zeleneev (2022): "Robust Estimation and Inference in Panels with Interactive Fixed Effects," *arXiv preprint arXiv:2210.06639*.

Bonhomme, S. (2021): "Teams: Heterogeneity, Sorting, and Complementarity," *arXiv preprint arXiv:2102.01802*.

Bonhomme, S. and M. Weidner (2022): "Minimizing sensitivity to model misspecification," *Quantitative Economics*, 13, 907–954.

Boschini, A. and A. Sjögren (2007): "Is team formation gender neutral? Evidence from coauthorship patterns," *Journal of Labor Economics*, 25, 325–365.

Bugni, F. A., I. A. Canay, and P. Guggenberger (2012): "Distortions of asymptotic confidence size in locally misspecified moment inequality models," *Econometrica*, 80, 1741–1768.

Bugni, F. A. and T. Ura (2019): "Inference in dynamic discrete choice problems under local misspecification," *Quantitative Economics*, 10, 67–103.

Chen, X., E. T. Tamer, and A. Torgovitsky (2011): "Sensitivity analysis in semiparametric likelihood models," *Cowles Foundation discussion paper*.

Christensen, T. and B. Connault (2023): "Counterfactual sensitivity and robustness," *Econometrica*, 91, 263–298.

Colussi, T. (2018): "Social ties in academia: A friend is a treasure," *Review of Economics and Statistics*, 100, 45–50.

Conley, T. G., C. B. Hansen, and P. E. Rossi (2012): "Plausibly exogenous," *Review of Economics and Statistics*, 94, 260–272.

Davezies, L., X. D'Haultfœuille, and Y. Guyonvarch (2021): "Empirical Process Results for Exchangeable Arrays," *The Annals of Statistics*, 49, 845–862.

Donoho, D. L. (1994): "Statistical estimation and optimal recovery," *The Annals of Statistics*, 22, 238–270.

DUCTOR, L. (2015): "Does co-authorship lead to higher academic productivity?" *Oxford Bulletin of Economics and Statistics*, 77, 385–407.

DUCTOR, L., M. FAFCHAMPS, S. GOYAL, AND M. J. VAN DER LEIJ (2014): "Social networks and research output," *Review of Economics and Statistics*, 96, 936–948.

FAFCHAMPS, M., M. J. VAN DER LEIJ, AND S. GOYAL (2010): "Matching and network effects," *Journal of the European Economic Association*, 8, 203–231.

GOYAL, S., M. J. VAN DER LEIJ, AND J. L. MORAGA-GONZÁLEZ (2006): "Economics: An emerging small world," *Journal of Political Economy*, 114, 403–412.

GRAHAM, B. (2020): "Handbook of Econometrics, volume 7A, chapter Network data," .

GRAHAM, B. S. (2022): "Sparse network asymptotics for logistic regression," *arXiv preprint arXiv.2010.04703*.

HSIEH, C.-S., M. KONIG, X. LIU, AND C. ZIMMERMANN (2022): "Collaboration in bipartite networks," Tech. rep., Working Paper.

HUBER, P. (1964): "Robust estimation of a location parameter," *Ann. Math. Statist.*, 35, 73–101.

HUBER, P. J. AND E. M. RONCHETTI (2009): *Robust statistics*, John Wiley & Sons Hoboken, NJ, USA.

IMBENS, G. W. (2003): "Sensitivity to exogeneity assumptions in program evaluation," *American Economic Review*, 93, 126–132.

KITAMURA, Y., T. OTSU, AND K. EVDOKIMOV (2013): "Robustness, infinitesimal neighborhoods, and moment restrictions," *Econometrica*, 81, 1185–1201.

LEAMER, E. E. (1985): "Sensitivity analyses would help," *The American Economic Review*, 75, 308–313.

LEUNG, M. (2015): "Two-step estimation of network-formation models with incomplete information," *Journal of Econometrics*, 188, 182–195.

MASTEN, M. A. AND A. POIRIER (2020): "Inference on breakdown frontiers," *Quantitative Economics*, 11, 41–111.

——— (2021): "Salvaging falsified instrumental variable models," *Econometrica*, 89, 1449–1469.

MENZEL, K. (2021): "Bootstrap with cluster-dependence in two or more dimensions," *Econometrica*, 89, 2143–2188.

NEVO, A. AND A. M. ROSEN (2012): "Identification with imperfect instruments," *Review of Economics and Statistics*, 94, 659–671.

NEWEY, W. K. (1985): "Generalized method of moments specification testing," *Journal of Econometrics*, 29, 229–256.

——— (1990): "Semiparametric efficiency bounds," *Journal of Applied Econometrics*, 5, 99–135.

NEWMAN, M. E. (2001): "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, 98, 404–409.

Ridder, G. and S. Sheng (2020): "Estimation of large network formation games," Working paper.

Rieder, H. (1994): *Robust Asymptotic Statistics*, Springer Verlag, New York, NY.

Tabord-Meehan, M. (2019): "Inference with dyadic data: Asymptotic behavior of the dyadic-robust t-statistic," *Journal of Business & Economic Statistics*, 37, 671–680.