A Nonparametric Test for Cross-Unit Spillovers*

Margherita Comola† Camila Comunello‡ Abhimanyu Gupta§

April 9, 2026

WORK IN PROGRESS

Abstract

We propose a novel nonparametric test for cross-unit spillovers that may operate through peers' attributes, peers' outcomes, or both. The test is straightforward to implement, as it requires only estimation under the null of no spillovers, and it is shown to have a convenient asymptotic standard normal distribution. It is also versatile, accommodating data generated by a wide range of interaction structures. We present three empirical illustrations showing that the test is effective at detecting cross-unit spillovers arising in a nonparametric manner that existing approaches may fail to uncover.

Keywords: Nonparametric test; Cross-unit spillovers; Social interactions; Interference

JEL Codes: C21, C14

* *Acknowledgements*: We are grateful to Marco Alfano, Aislinn Bohren, Arun Chandrasekhar, Ben Deaner, Áureo de Paula and Michael Vlassopoulos for helpful comments. Comola acknowledges the financial support of the French National Research Agency (ANR-21-CE26-0002-01). Comunello and Gupta acknowledge the financial support of the Leverhulme Trust via grant RPG-2024-038.

† University Paris-Saclay (RITM) and Paris School of Economics. Email: margherita.comola@psemail.eu

‡ University of Essex. Email: c.comunello@essex.ac.uk

§ Queen's University and University of Essex. Email: abhimanyu.g@queensu.ca

1 Introduction

Cross-unit dependence pervades many empirical settings and poses a fundamental challenge for econometric inference. This paper deals with the two most common sources of cross-unit spillovers, which arise through the attributes channel and the outcome channel, respectively. To illustrate, consider the canonical cross-sectional linear model, where the outcome of the unit (e.g. individual) y_i is regressed on own attribute(s) x_i. First, the attributes of other units ('peers') may affect i's outcome. We term this as **'covariate (c) spillovers'**. The second source of cross-unit dependence refers to the case where the outcomes of other units affect i's outcome. We call this **'outcome (y) spillovers'**. In this paper, we propose a test for unknown nonparametric spillovers operating through one or both channels, establish its asymptotic properties, and illustrate its applicability in three diverse settings.

Cross-unit spillovers have received considerable attention from applied economists in a broad range of contexts. These include, *inter alia*, disease transmission (Miguel and Kremer, 2004; Ozier, 2018), educational outcomes (Sacerdote, 2001; Lalive and Cattaneo, 2009; Bobonis and Finan, 2009; Avvisati, Gurgand, Guyon, and Maurin, 2013), employment decisions (Duflo and Saez, 2003; Brown and Laschever, 2012) and technology adoption (Oster and Thornton, 2012; Banerjee, Chandrasekhar, Duflo, and Jackson, 2013; Cai, De Janvry, and Sadoulet, 2015). A first strand of the literature on (broadly defined) 'peer effects' has explicitly modeled cross-unit dependence in observational data via the attribute and/or the outcome channel, depending on the setting. Oftentimes, cross-unit dependence is modeled solely through the attribute channel, even though outcome spillovers could also be incorporated due to economic considerations.[1] A related line of work focuses on treatment-mediated spillovers, which are a first-order concern in the context of impact evaluation as they violate the Stable Unit Treatment Value Assumption (SUTVA), which asserts that an individual's potential outcomes should be independent of peers' treatment assignments.[2] In response to this concern, it has become increasingly common to design cluster-randomized experiments generating exogenous variation in peers' treatment status.[3]

[1] The exclusion restriction that peers' attributes serve as a reduced-form sufficient statistic for their outcomes is frequently imposed. However, when the research design permits, spillovers operating through both peers' covariates and peers' realized outcomes can be jointly identified, offering a sharper understanding of the underlying economic mechanisms (Bramoullé, Djebbari, and Fortin, 2009; Bursztyn and Fiorin, 2017).

[2] For simplicity, in this paper we abstract from the issue of contamination, whereby subjects in a randomized trial may move from the treatment group to the control group.

[3] Cluster-randomized trials (also known as 'two-stage randomization experiments', 'randomized-

We develop a nonparametric test for cross-unit dependence arising via either the attribute or outcome channels, or both. The idea behind our test is reminiscent of classical Lagrange Multiplier (LM)-type diagnostic tests, such as the RESET test. We construct a test statistic to assess cross-unit dependence against the alternative that such features may arise in a nonparametric manner, which is operationalized by approximating the nonparametric functions of the alternative with a series of underlying basis functions. However, because we apply an LM-type approach, we only need to estimate the model under the null of no spillovers. Thus, our test has the feature of being nonparametric but requiring only standard parametric estimates of a familiar multiple linear regression. Asymptotic theory is provided, showing that our test has an asymptotic standard normal distribution under the null and is consistent in the sense of having asymptotically unit probability of rejecting a false null. We derive these properties under a cluster-robust framework, thus allowing the incorporation of what is now common practice in applied work. Extensions to alternative error dependence structures—such as serial correlation or more general spatial dependence—are conceptually straightforward.

Failure to account for nonlinearities may undermine the reliability of empirical findings. This concern is compounded by the fact that spillovers are typically modeled in a linear fashion, namely by including a linear function of peers' $\boldsymbol{c}$ and/or $\boldsymbol{y}$ as a regressor. Yet, in many economically relevant settings, agents' strategic behavior may depend on the entire distribution of peers' attributes and outcomes. In such contexts, cross-unit dependence is likely to generate nonlinear simultaneous determination of outcomes that linear specifications may fail to capture adequately. For example, when only a small fraction of peers adopt a given behavior or technology, agents may find it optimal to engage in complementary actions, whereas a high prevalence of adopters among peers may instead induce substitutability in best-response behavior.[4] If the estimates at different points of the distribution offset or dilute, a linear specification may not detect cross-unit spillovers appropriately, as our test does.

Our test is far-reaching in that it is versatile in its data requirements, which is an important advantage. First, it accommodates data defined through a variety of interaction structures, including those based on blocks or links. Block-type data are partitioned

saturation experiments', 'partial-population experiments') randomly assign different treatment rates across different clusters.

[4]Nonlinearity in the social adoption rate was documented both theoretically and empirically (Bandiera and Rasul, 2006; Young, 2009; Acemoglu, Özdaglar, and Tahbaz-Salehi, 2016).

into separate self-exclusive groups within which all units are assumed to interact. Blocks may represent villages or schools or nuclear households for individuals, geographical area and/or productive sectors for firms. Alternatively, network-type data contain detailed links between units, which may or may not overlap (e.g. i is linked to j and j is linked to k, but i is not linked to k). This is the case for self-declared link data in household surveys, or trade data among firms from administrative records. Our test accommodates both data structures. Second, it allows for heterogeneous spillovers via multiple interaction matrices, as we justify in Appendix B.1. Third, it allows the interaction structure to be incomplete or measured noisily. In applied work, interaction data are often measured poorly yet still convey useful information. By embedding our test within a latent-space (embedded-graph) framework, Appendix B.2 derives conditions under which the perturbations induced by measurement error are asymptotically negligible, ensuring that inference remains valid.

While this paper proposes a nonparametric test for cross-unit spillovers, most of the methodological literature on cross-unit dependence has focused on the estimation of causal parameters under interference. By and large, the existing approaches rely on parametric assumptions to achieve this aim (Rosenbaum, 2007; Hudgens and Halloran, 2008; Hirano and Hahn, 2010; Liu and Hudgens, 2014; Baird, Bohren, McIntosh, and Özler, 2018; Arduini, Patacchini, and Rainone, 2020; Viviano, Lei, Imbens, Karrer, Schrijvers, and Shi, 2025; McNealis, Moodie, and Dean, 2024). One exception is Vazquez-Bare (2023), who works in a nonparametric identification framework to propose estimators for spillover effects in experiments. Thus, we view this paper as complementing this literature by providing applied practitioners with an easy-to-implement diagnostic tool to guide the design, validation, and refinement of estimation strategies. If no spillovers are detected, standard estimation strategies—such as linear intent-to-treat regressions—are justified. If instead the test detects nonlinear cross-unit dependence, the estimation strategy should be adapted to account for its potentially substantial effects.[5] Our test can also be useful prior to rolling out a large-scale survey, when researchers, based on the pilot, must decide whether and how to adjust the design to account for cross-unit dependence.[6]

[5]While our test does not explicitly suggest the nonparametric functional form which best describes the data at hand, a variety of suitable econometric methods are available, see e.g. Jenish (2012, 2016); Xu and Lee (2015, 2018).

[6]For instance, a researcher may apply our test to interaction data collected in a pilot study to assess whether detailed link information is required, or whether the treatment intensity should be exogenously varied across clusters in a subsequent scale-up. Because both design choices can entail substantial additional survey costs, they are warranted primarily when spillovers are expected to play an important role.

We illustrate the implementation, scope, and economic relevance of our test through three empirical applications that revisit studies of spillovers in very different settings. In the first illustration, we examine peer effects in a high skill environment, namely professional golf tournaments (Guryan, Kroft, and Notowidigdo, 2009). The second illustration revisits the study of network frictions in small firms' performance in China by Cai and Szeidl (2017). The final example examines a field experiment in the Netherlands to study student achievement in response to changes in the academic environment (Booij, Leuven, and Oosterbeek, 2017). Across all three settings, we show that our test is able to detect cross-unit dependence through peers' attributes and/or peers' outcomes in some —but not all— cases where linear functional forms fail to do so.

This paper is organized as follows. Section 2 introduces the test statistic, while Section 3 characterizes its asymptotic behavior. In Section 4 we illustrate the test by revisiting three existing studies. Section 5 concludes. We show that our test controls size well in a Monte Carlo study of finite performance in Appendix A. In Appendix B we extend our method to heterogeneous cross-unit dependence, and to embedded graphs with noisy measurement and parametric modeling of the underlying link formation structure. Appendices C and D contain theorem proofs and auxiliary lemmas respectively, while Appendix E reproduces the original results for our three empirical demonstrations.

2 Method and test statistic

We write the general model with both $\boldsymbol{c}$ and $\boldsymbol{y}$ spillovers in scalar notation as

$$y_i = f(w_i'y) + x_i'\beta + \sum_{j=1}^{l} g_j(w_i'c_j) + \epsilon_i, \quad i = 1, ..., n, \tag{2.1}$$

where $f(\cdot) : \mathbb{R} \to \mathbb{R}$ is an unknown function that captures 'outcome ($\boldsymbol{y}$) spillovers', and $g_j(\cdot) : \mathbb{R} \to \mathbb{R}$ are also unknown functions that capture 'attribute ($\boldsymbol{c}$) spillovers'. $W = (w_1, \ldots, w_n)'$ is a social weight matrix representing cross-unit interactions that is either fixed or exogenous conditional on observables, and has zero diagonal.[7] x_i' is the i-th row

[7]Since our approach is cast within an instrumental-variables framework, it naturally accommodates endogeneity of the term $w_i'y$ whenever valid instruments are available. In the context of network-type data, it is compatible with instrumental-variable strategies designed to address outcome spillovers that are endogenous due to simultaneity (Bramoullé et al., 2009) or due to assortative link formation on unobservables (Jochmans, 2023).

of an $n \times k$ regressor matrix X that can have endogenous elements as long as instruments are available, c_j, $j = 1, \ldots, l$, are some $n \times 1$ vectors of exogenous regressors and ϵ_i is an unobserved disturbance. The data are observed as n_g observations in each of $g = 1, \ldots, G$ clusters, so that $n = \sum_{g=1}^{G} n_G$. We take n_g as fixed with $\sup_{g=1,\ldots,G} n_g < \infty$ and therefore $n \sim G$, i.e. our sample size grows like the number of clusters. Our approach gives rise to three different testing options:

1. Jointly test for both $\boldsymbol{c}$ and $\boldsymbol{y}$ spillovers: We term this the '$\boldsymbol{cy}$ test'.
2. Omit $f(w_i'y)$ from the general model (2.1) and test for only $\boldsymbol{c}$ spillovers: We term this the '$\boldsymbol{c}$ test'.
3. Omit $g_j(s) = 0,\ j = 1, \ldots, l$, from the general model (2.1) and test for only $\boldsymbol{y}$ spillovers: We term this the '$\boldsymbol{y}$ test'.

The corresponding null hypotheses are:

$$\boldsymbol{cy} \text{ test, } \mathcal{H}_0 : f(s) = 0 \text{ and } g_j(s) = 0,\ j = 1, \ldots, l, \tag{2.2}$$

$$\boldsymbol{c} \text{ test, } \mathcal{H}_0 : g_j(s) = 0,\ j = 1, \ldots, l, \tag{2.3}$$

$$\boldsymbol{y} \text{ test, } \mathcal{H}_0 : f(s) = 0, \tag{2.4}$$

for all $s \in support(s)$, which implies in all cases that the null model is $y_i = x_i'\beta + \epsilon_i$ i.e. a standard linear regression.

The $\boldsymbol{cy}$ test jointly includes nonlinear spillovers in both the covariate/attribute and outcome channels, while the $\boldsymbol{c}$ test and $\boldsymbol{y}$ test examine the covariate and outcome channels individually, respectively. In our applications, we implement the $\boldsymbol{c}$ and $\boldsymbol{cy}$ tests. Most applied work focuses on linear spillovers in covariates and extending that to nonlinearity via the $\boldsymbol{c}$ test seems natural, while augmenting for outcome spillovers through the $\boldsymbol{cy}$ test demonstrates the full power of our approach. For ease of exposition we present asymptotic results and notation for the most general case covered by the $\boldsymbol{cy}$ test in (2.2).

Let $\psi_i(s)$, for $i = 1, \ldots, p$, be a user-chosen set of basis functions (our applications use Hermite polynomials) such that

$$f(s) = \sum_{i=1}^{p} \mu_{f,i}\psi_i(s) + r_f(s), \quad g_j(s) = \sum_{i=1}^{p} \mu_{g_j,i}\psi_i(s) + r_{g_j}(s),\ j = 1, \ldots, l, \tag{2.5}$$

with $p = p_n$ a divergent deterministic sequence i.e. $p \to \infty$ as $n \to \infty$, $\mu_f = (\mu_{f,1}, \ldots, \mu_{f,p})'$, $\mu_{g_j} = (\mu_{g_j,1}, \ldots, \mu_{g_j,p})'$ vectors of unknown series coefficients and $r_f(s)$, $r_{g_j}(s)$ approximation errors. We define our approximate null hypothesis as

$$\mathcal{H}_{0A} : \mu_f = 0 \text{ and } \mu_{g_j} = 0,\ j = 1, \ldots, l,\ \text{for some } \beta, \tag{2.6}$$

which is a set of $q = p(l+1)$ restrictions, so that $q \to \infty$ as $n \to \infty$. This indicates that standard fixed-dimension asymptotic will not work for our test. Our test statistic is based on determining if the moment conditions for the instrumental variables (IV) estimate of β under the null hypothesis are close enough to zero. OLS is obviously a special case of this.

Now, for each $i = 1, \ldots, p$, define the $n \times 1$ vector $\Upsilon_{f,i}(y) = (\psi_i(w_1'y), \ldots, \psi_i(w_n'y))'$ and the $n \times 1$ vectors $\Upsilon_{g_j,i}(c_j) = (\psi_i(w_1'c_j), \ldots, \psi_i(w_n'c_j))'$, and write

$$\Upsilon_{f,g} = \begin{pmatrix} \Upsilon_{f,1}(y) & \ldots & \Upsilon_{f,p}(y) & \Upsilon_{g_1,1}(c_1) & \ldots & \Upsilon_{g_1,p}(c_1) & \ldots & \Upsilon_{g_l,1}(c_l) & \ldots & \Upsilon_{g_l,p}(c_l) \end{pmatrix},$$

which is an $n \times q$ matrix. Denote $\mu = (\mu_f, \mu_{g_1}, \ldots, \mu_{g_l})'$. Given the expansion in (2.5), the series approximated IV objective function is

$$\mathcal{F}_p(\beta, \mu, y) = \frac{1}{n} \left(y - \Upsilon_{f,g}\mu - X\beta\right)' \mathcal{P}_Z \left(y - \Upsilon_{f,g}\mu - X\beta\right), \tag{2.7}$$

where Z is an $n \times m$ matrix of valid instruments, with $m \geq q + k$, and $\mathcal{P}_Z = Z(Z'Z)^{-1}Z'$. Next, define the $n \times (q+k)$ matrix $U = \begin{pmatrix} \Upsilon_{f,g_j} & X \end{pmatrix}$. OLS is a special case with $Z = U$.

Define the $(q+k) \times 1$ gradient vector $\tilde{d}(\beta, y)$ of (2.7) under $\mathcal{H}_{0A}$ as

$$\tilde{d}(\beta, y) = \frac{\partial \mathcal{F}(\mu, \beta, y)}{\partial (\mu, \beta)'} = -\frac{2}{n} U' \mathcal{P}_Z (y - X\beta). \tag{2.8}$$

Denoting by $\hat{\beta}$ some consistent estimate of β, e.g. IV or OLS, under $\mathcal{H}_{0A}$, the gradient evaluated at the corresponding residuals is

$$\hat{d} = \tilde{d}\left(\hat{\beta}, y\right) = -\frac{2}{n} U' \mathcal{P}_Z (y - X\hat{\beta}). \tag{2.9}$$

Let $\hat{J} = n^{-1}Z'U$, where $\hat{J}$ is $m \times (q+k)$. Next, define the $m \times m$ matrices $\hat{M} = n^{-1}Z'Z$ and $\hat{\Phi} = n^{-1}Z'\hat{\Sigma}Z$, with $\hat{\Sigma} = diag\left(\hat{\Sigma}_1, \ldots, \hat{\Sigma}_G\right)$ where $\hat{\Sigma}_g$ has typical (i,j)-th element $\hat{\epsilon}_i \hat{\epsilon}_j$ and $\hat{\epsilon}_i = y_i - x_i'\hat{\beta}$, for $i, j = 1, \ldots, n_g$ and $g = 1, \ldots, G$. Thus the covariance matrix of the gradient evaluated at the estimates is

$$\hat{H} = 4\hat{J}'\hat{M}^{-1}\hat{\Phi}\hat{M}^{-1}\hat{J}, \tag{2.10}$$

and we hence define our cluster robust test statistic as

$$\mathcal{S} = \frac{n\hat{d}'\hat{H}^{-1}\hat{d} - q}{\sqrt{2q}}. \tag{2.11}$$

This is a weighted measure of the distance of the gradient from zero, centred and rescaled to account for $q \to \infty$. In practice, especially for small q, one can use $n\hat{d}'\hat{H}^{-1}\hat{d}$ as the test statistic with χ^2_q critical values instead of standard normal ones. We study this in our simulations in Appendix A.

3 Asymptotic theory

We commence this section by introducing some technical assumptions to establish the limiting behaviour of (2.11) under $\mathcal{H}_{0A}$. Our approach is based on the series based tests in Gupta (2018) and Gupta, Lee, and Rossi (2025). Throughout we denote by K a generic positive constant, arbitrarily large but independent of p and n.

Assumption 1. *ϵ_i are random variables with zero mean and unknown variance $\sigma_i^2 \in [c, K]$, $c > 0$, and, for some $\tau > 0$, $\mathbb{E}\left|\epsilon_i\right|^{8+\tau} \leq K$ for $i = 1, \ldots, n$. Furthermore ϵ_{ig} and $\epsilon_{jg'}$ are independent for $g \neq g'$, $g, g' = 1, \ldots, G$, while $\mathbb{E}\left(\epsilon_{ig}\epsilon_{jg}\right) = \sigma_{ijg} < \infty$, $i \neq j$, and we accordingly write $\Sigma = diag\left(\Sigma_1, \ldots, \Sigma_G\right)$.*

Assumption 2. *$\mathbb{E}(x_{ir}^4) \leq K$ and $\mathbb{E}(z_{is}^4) \leq K$, for $i = 1, \ldots, n$ and $r = 1, \ldots, k$ and $s = 1, \ldots, l$.*

We also allow $cov(\epsilon_i, x_{ij}) \neq 0$, for some $j = 1, \ldots, k$, i.e. X might contain some endogenous columns. Let X_1 be the $n \times k_1$ matrix containing the subset of exogenous columns of X, while X_2 ($n \times k_2$, with $k_2 = k - k_1$) contains the endogenous ones. Now, for a generic symmetric positive-definite matrix A, let $\overline{eig}(A)$ and $\underline{eig}(A)$ denote its largest and smallest eigenvalues, respectively. For a generic matrix B, denote by $\|B\| = \sqrt{\overline{eig}(B'B)}$, i.e. the spectral norm of B, and by $\|B\|_\infty$ its largest absolute row sum.

Assumption 3. *The $n \times n$ matrix Σ satisfies*

$$\limsup_{n\to\infty} \sup_{g=1,\ldots,G} \overline{eig}(\Sigma_g) < \infty, \quad \liminf_{n\to\infty} \inf_{g=1,\ldots,G} \underline{eig}\left(\Sigma_g\right) > 0, \tag{3.1}$$

the $m \times m$ *matrix* $M = \mathbb{E}(\hat{M})$, *with* $m \geq q + k$, *satisfies*

$$\limsup_{n\to\infty} \overline{eig}(M) < \infty, \quad \liminf_{n\to\infty} \underline{eig}(M) > 0, \tag{3.2}$$

and the $(q + k) \times (q + k)$ *matrix* $L = n^{-1}\mathbb{E}(U'U)$ *satisfies*

$$\limsup_{n\to\infty} \overline{eig}(L) < \infty, \quad \liminf_{n\to\infty} \underline{eig}(L) > 0, \tag{3.3}$$

for n *large enough. For some* $\nu > 0$ *satisfying* $n/p^{(\nu+1/2)} = o(1)$,

$$\sup_z r_f(z) + \sup_{j=1,\ldots,l} \sup_z r_{g_j}(z) = O_p\left(p^{-\nu}\right),$$

as $p \to \infty$. $\mathbb{E}\left(u_{il_2}^4\right) \leq K$ *for* $i = 1, \ldots, n, l_1 = 1, \ldots, m$ *and* $l_2 = 1, \ldots, q + k$, *and* ϵ_i *and* z_j *are uncorrelated for each* $i, j = 1, \ldots, n$.

Assumption 3 imposes regularity conditions and controls the approximation errors. Specifically, (3.2)-(3.3) are asymptotic boundedness and no multicollinearity conditions for matrices of increasing dimension, while under Assumption 1, (3.1) also ensures that $0 < \sup_{i=1,\ldots,n} \Sigma_i < \infty$ in the special case of purely heteroskedasticity robust testing i.e. $G = n$ and Σ_i are scalars. For the instruments Z we use at least k_2 columns of instruments for the endogenous covariates X_2, and also the columns of X_1, WX_1. We also use a set of instruments of the form $\psi_r\left(\sum_j w_{ij}x_{1,jl}\right)$, where $r = 1, \ldots, p$, and $x_{1,jl}$ denotes the (j, l)th element of X_1, with $l = 1, \ldots, k_1$. For more discussion on approximation error decay rates see e.g. Chen (2007). Our next assumption sets a suitable bound on cross-sectional dependence, analogous to that in Lee and Robinson (2016). Conditions such as linear process representations for the underlying random variables or the near epoch dependence conditions of Jenish and Prucha (2012) imply that this assumption holds.

Assumption 4. *Let*

$$\xi = \sup_{0\leq l,k\leq m} \left(\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \mathbb{E}(z_{il}z_{ik}z_{jl}z_{jk}) - \left(\mathbb{E}\left(\sum_{i=1}^{n} z_{il}z_{ik}\right)\right)^2 \right)$$

$$\varkappa \;=\; \sup_{0\le l\le m,\;0\le k\le q+k}\left(\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\mathbb{E}(z_{il}u_{ik}z_{jl}u_{jk})-\left(\mathbb{E}\left(\sum_{i=1}^{n}z_{il}u_{ik}\right)\right)^{2}\right),$$

and assume

$$\xi+\varkappa=O(n)\quad as\quad n\to\infty. \tag{3.4}$$

Our null asymptotic theory first approximates the test statistic $\mathcal{S}$ with a quadratic form in ϵ, and then shows that this approximation is asymptotically standard normal. Write $J=E(\hat{J})$ and define

$$\begin{aligned} d=d(\beta_0,y) &= -\frac{2}{n}J'M^{-1/2}\left(I-M^{-1/2}N\left(N'M^{-1}N\right)^{-1}N'M^{-1/2}\right)M^{-1/2}Z'\epsilon \\ &= -\frac{2}{n}J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}Z'\epsilon, \end{aligned} \tag{3.5}$$

where $\mathcal{K}_{NM}=\left(I-M^{-1/2}N\left(N'M^{-1}N\right)^{-1}N'M^{-1/2}\right)$ is $m\times m$ and $N=\mathbb{E}(\hat{N})$, with $\hat{N}=n^{-1}Z'X$, the last being an $m\times k$ matrix with full rank under (3.3) in Assumption 3. Set

$$H=n\mathbb{E}(dd')=4J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}\Phi M^{-1/2}\mathcal{K}_{NM}M^{-1/2}J, \tag{3.6}$$

with $\Phi=n^{-1}\mathbb{E}(Z'\Sigma Z)$. Under Assumptions 1 and 3, H^{-1} exists and is non-singular for n large enough, using the following lemma for the eigenvalues of Φ.

We now state the main result of this section.

Theorem 1. *Under $\mathcal{H}_0$, Assumptions 1-4, $\nu>5/2$, and $p^3/n=o(1)$,*

$$\mathcal{S}\xrightarrow{d}N(0,1),\ \ as\ n\to\infty. \tag{3.7}$$

Theorem 1 provides asymptotic justification for using one-sided, standard normal critical values as observed also by Hong and White (1995). We also provide guidance for an empirical choice of $p=\left[n^{1/3}\right]$, where $[\cdot]$ denotes the closest integer. However, note that the rate $p^3/n\to$ is equivalent asymptotically to $p^3(l+1)/n\to$ because l is fixed, but in finite samples the extra $l+1$ factor can play a role. Thus our recommendation is to use[8]

$$\text{For the } \boldsymbol{cy} \text{ test, with both } \boldsymbol{c} \text{ and } \boldsymbol{y} \text{ spillovers} \quad : \quad p_{cy} = \frac{\left\lceil n^{1/3} \right\rceil}{l+1}, \tag{3.8}$$

$$\text{For the } \boldsymbol{c} \text{ test, with only } \boldsymbol{c} \text{ spillovers} \quad : \quad p_{c} = \frac{\left\lceil n^{1/3} \right\rceil}{l}, \tag{3.9}$$

$$\text{For the } \boldsymbol{y} \text{ test, with only } \boldsymbol{y} \text{ spillovers} \quad : \quad p_{y} = \left\lceil n^{1/3} \right\rceil. \tag{3.10}$$

Our next theorem relates to the power properties of our test. First consider the global alternative

$$\mathcal{H}_{1A}: \quad \mu_i \neq 0, \quad \text{for some } i = 1, \ldots, q, \text{ and any } \beta, \tag{3.11}$$

where μ_i denotes the i-th element of μ. We introduce the unrestricted quantities

$$\epsilon_{Ui}(\mu, \beta) = y_i - \mu' \upsilon_{f,g,i} - \beta' x_i, i = 1, \ldots, n, \quad \text{and} \quad \tilde{\Phi}_U = \tilde{\Phi}_U(\mu, \beta) = n^{-1} Z' \tilde{\Sigma}_U Z, \tag{3.12}$$

where $\upsilon'_{f,g,i} = (\psi_1(w_i'y), \ldots, \psi_p(w_i'y), \psi_1(w_i'c_1), \ldots, \psi_p(w_i'c_1), \ldots, \psi_1(w_i'c_l), \ldots, \psi_p(w_i'c_l))'$, $\tilde{\Sigma}_U = diag\left(\tilde{\Sigma}_{U1}, \ldots, \tilde{\Sigma}_{UG}\right)$, and $\tilde{\Sigma}_{Ug}$ has (i,j)-th element $\epsilon_{Ui}(\mu, \beta)\epsilon_{Uj}(\mu, \beta)$. Then $\hat{\Phi} = \tilde{\Phi}_U(0_{q \times 1}, \hat{\beta})$. Let $\gamma = (\mu, \beta) \in \Gamma = \Re^q \times \Re^k$ and introduce:

Assumption 5. *For all sufficiently large n and all $j = 1, \ldots, q+k$,*

$$\sup_{\gamma \in \Gamma} \overline{eig}(\tilde{\Phi}_U) + \sup_{\gamma \in \Gamma} \overline{eig}\left(\frac{\partial \tilde{\Phi}_U}{\partial \gamma_j}\right) = O_p(1), \tag{3.13}$$

and

$$\left\{\inf_{\gamma \in \Gamma} \underline{eig}(\tilde{\Phi}_U)\right\}^{-1} + \left\{\inf_{\gamma \in \Gamma} \underline{eig}\left(\frac{\partial \tilde{\Phi}_U(\gamma)}{\partial \gamma_j}\right)\right\}^{-1} = O_p(1). \tag{3.14}$$

This assumption imposes mild regularity under $\mathcal{H}_{1A}$, reminiscent of boundedness and invertibility conditions. Our power result follows below.

Theorem 2. *Under Assumptions 1-5, $\mathcal{H}_{1A}$, $\nu > 5/2$, and $p^3/n = o(1)$, $\mathcal{S}$ provides a consistent test.*

[8] In Appendix B.1 we also show that we can extend our test to a setting with multiple W matrices, say ℓ, for which we recommend $p_{cy,m} = \frac{\left\lceil n^{1/3} \right\rceil}{\ell + l}$ in (B.5) therein.

4 Three demonstrations

In this section, we illustrate the scope of our nonparametric test by revisiting three existing studies that investigate peer effects across different contexts (high-skill professionals, firms, and students' performance) with mixed findings. We show that our test is able to detect spillovers in some cases where linear specifications fail to do so. We use Hermite polynomials as basis functions in all three examples. Additional details on the original studies are reported in Appendix E.

4.1 Professional golf tournaments (Guryan et al., 2009)

4.1.1 Context and findings

Our first example builds on Guryan et al. (2009), who study whether peer effects influence individual productivity in high-skill professional environments with an application to professional golf tournaments. They exploit a natural experiment within golf tournaments where playing partners are randomly assigned within predefined block–round categories.

Their original results are reproduced in Appendix E, Table E1 and include three specifications.[9] In the specification (i), players' performance is modeled as a function of their own ability, measured by the corrected handicap score, and the average ability of their peers in the same block–round–tournament.[10] Specification (ii) incorporates alternative measures of peer ability—average driving distance, number of putts, and number of greens hit—designed to distinguish motivation effects (e.g., higher effort induced by stronger partners) from learning effects (e.g. adapting to observed putting strategies). Specification (iii) introduces heterogeneity by interacting partners' average ability with a player's own baseline ability and years of professional experience.

While previous studies have found significant positive peer effects in low-skill labour markets (Bandiera, Barankay, and Rasul, 2009; Mas and Moretti, 2009), Guryan et al. (2009) find limited evidence of peer effects in individual performance. In particular, they conclude against peer effects in specifications (i) and (ii), while specification (iii) provides some support for heterogeneous peer effects via the experience channel.

[9] Because of the large sample size and relatively few covariates, this particular example leads to large choices for p which can cause some rank deficiency in finite samples. We deal with this by dropping nearly collinear columns of Z and adding new ones using transformations of the regressors until the rank condition is fulfilled.

[10] For details about the way the corrected handicap score is calculated, see Appendix E.

4.1.2 Test results

Table 1 summarizes the authors' main results and the results from our tests. The three rows correspond to the three original specifications reported in Appendix E, Table E1, in order of appearance. Columns 1-3 report the specification number, the construction of attribute peer exposure variables $w'c$ and the number of peer attribute terms (l). The "Original result" column reports the conclusions reached in the original paper for the corresponding null hypothesis of no peer effects. Recall that the $\boldsymbol{c}$ test examines dependence operating through peers' attributes, while the $\boldsymbol{cy}$ test augments the instrument set with functions of peers' outcomes. This latter test goes beyond the original specifications by allowing for nonparametric propagation through contemporaneous peer performance as well. The columns p_c and p_{cy} report the number of basis functions used in each test and n reports the sample size.

Our findings support the evidence of peer effects in professional golfing in ways that the linear specifications in Guryan et al. (2009) do not capture. When we assume spillovers from attributes only ($\boldsymbol{c}$ test), contrary to the authors' findings we detect spillovers when ability is measured with alterative measures, as in specification (ii). Additionally, our $\boldsymbol{cy}$ test also finds evidence of spillovers in contemporaneous peer scores of specification (i) - an effect also undetected in the original findings.

4.2 Network expansion and firm performance (Cai and Szeidl, 2017)

4.2.1 Context and findings

Firms in developing economies face not only financial and managerial constraints but also networking frictions—such as limited trust or information—that may prevent them from accessing knowledge, clients, and suppliers.[11] In our second illustration we revisit the study of Cai and Szeidl (2017), who investigate whether an exogenous expansion in business networks can improve firm performance.

The intervention under study randomly assigned firms into small business-association groups of ten owner-managers. The treatment group managers met monthly for one year, while other firms' managers did not participate and served as a control group. The intervention outcomes were measured through detailed baseline, midline, and endline surveys

[11] Such networking frictions are likely to be more binding in developing economies, where search and trust costs impede self-organization, whereas at higher levels of development similar business associations can often emerge without external coordination (Cai and Szeidl, 2017).

Table 1: Peer Exposure Design and Test Results: Guryan et al. (2009)

Dependent variable: Score in a given tournament-round, $y_{i,tr}$

Peer exposure in outcomes: $w'_{i,tr}y_{tr}$

	Peer exposure in attributes					Peer exposure in outcomes		
	Peer exposure $w'c$	l	Original result	p_c	$\boldsymbol{c}$ test	p_{cy}	$\boldsymbol{cy}$ test	n
(i)	$w'_{i,tr}Ability$	1	Do not reject	26	Do not reject	13	Reject	17,492
(ii)	$w'_{i,tr}DrivDist$, $w'_{i,tr}Greens$, $w'_{i,tr}Putts$	3	Do not reject	9	Reject	6	Reject	17,182
(iii)	$w'_{i,tr}Ability$, $Ability_i \times w'_{i,tr}Ability$, $Exp_i \times w'_{i,tr}Ability$	3	Reject	9	Reject	6	Reject	17,492

Note: Each row reports results from a separate regression specification corresponding to models estimated in the original paper. Row (i) corresponds to the baseline model from Column 1 of Table 4 from Guryan et al. (2009), while rows (ii) and (iii) correspond to Column 5 of Table 5 and Column 4 of Table 8, respectively. Columns 1-3 report the specification number, the construction of attribute peer exposure variables and the number of peer attribute terms (l). The "Original result" column reports the conclusions reached in the original paper for the corresponding null hypothesis of no peer effects. The "$\boldsymbol{c}$ test" examines dependence operating through peers' attributes, while the "$\boldsymbol{cy}$ test" augments the instrument set with functions of peers' outcomes. The columns p_c and p_{cy} report the number of basis functions used in each test and n reports the sample size. The outcome variable is the golf score for the round. $Ability_i$ is measured by the player's average handicap, $DrivDist_i$ is the average driving distance, $Greens_i$ is the average number of greens hit in regulation, $Putts_i$ is the average number of putts per round, all averages over the previous 2-3 years. Exp_i experience is measured as years of experience. Peer exposure is measured using a weighted-average social vector, $w_{i,tr}$, where each player is exposed to the weighted average attribute/outcome of peers within the same group-tournament-round. Sample weights are given by the inverse of the sample variance of the estimated ability of each player, in line with the original study. All specification controls are identical to those reported in Table E1 in Appendix E.1. Tests are performed at the 95% confidence level. Standard errors are clustered at the playing group level.

covering sales, profits, employment, assets, inputs, management practices, and networks. The authors first examine the direct impact of the intervention and find that treated firms experienced significant and sustained improvements in performance.[12] To help uncover the mechanisms behind these gains, the authors then focus on peer composition within the networking groups and run a battery of peer-effect specifications that we focus on, reported in Table E2 of Appendix E.2. By proxying peer quality by baseline employment size, Cai and Szeidl (2017) show that firms randomly assigned to groups with larger peers achieved faster growth across multiple dimensions, including sales, profits, management practices, and network expansion.

4.2.2 Test results

Table 2 compares our nonparametric test results to the findings by Cai and Szeidl (2017). Each row (i)–(xiv) reports results from a separate regression corresponding to the specifications reported in Table E2. As before, for each of the outcomes we test for dependence operating through peers' attributes alone ($\boldsymbol{c}$ test) and through both peer attributes and outcomes ($\boldsymbol{cy}$ test). The columns p_c and p_{cy} report the number of basis functions used in each test, and n reports the sample size.

Table 2 reveals several differences between our nonparametric test results and the original findings. Our $\boldsymbol{c}$ test agrees with the authors' rejection of the null of no spillovers for sales, profits, utility costs, and management practices. However, we also reject the null for a number of specifications for which the authors find no significant peer effects. In particular, the $\boldsymbol{c}$ test detects spillovers in the number of employees, material costs, productivity, the number of suppliers, and innovation, whereas the original linear regressions report insignificant effects for these outcomes. Conversely, for total assets, bank loans, reported (book) sales, and the tax-to-sales ratio, both approaches fail to reject the null of no spillovers.

When nonparametric propagation through contemporaneous peer performance is permitted in the $\boldsymbol{cy}$ test, the null hypothesis is rejected for most outcomes. The only exceptions remain total assets, bank loans, reported sales, and tax-to-sales ratio. This pattern suggests that peer effects in this setting operate not only through peers' observable characteristics

[12] Sales increased by about 8% at midline and 10% at endline for treated firms relative to control firms, with similarly positive impacts on profits, employment, fixed assets, and input usage. Firms expanded both the number of clients and suppliers they interacted with, increased access to formal and informal borrowing, and improved management scores by roughly a fifth of a standard deviation.

Table 2: Peer Exposure Design and Test Results: Cai and Szeidl (2017)

Peer exposure in attributes: $w'_{i,t}EmpSize$

Peer exposure in outcomes: $w'_{i,t}y_t$

$l = 1$

		Peer exposure in attributes			Peer exposure in outcomes		
Spec.	y	Original result	p_c	$\boldsymbol{c}$ test	p_{cy}	$\boldsymbol{cy}$ test	n
(i)	Sales	Reject	16	Reject	8	Reject	4,183
(ii)	Profits	Reject	16	Reject	8	Reject	4,076
(iii)	No. of employees	Do not reject	16	Reject	8	Reject	4,183
(iv)	Total Assets	Do not reject	16	Do not reject	8	Do not reject	4,183
(v)	Material Cost	Do not reject	16	Reject	8	Reject	4,148
(vi)	Utility Cost	Reject	16	Reject	8	Reject	4,086
(vii)	Productivity	Do not reject	16	Reject	8	Reject	4,183
(viii)	No. of clients	Reject	16	Reject	8	Reject	4,173
(ix)	No. of Suppliers	Do not reject	16	Reject	8	Reject	4,170
(x)	Bank loan	Do not reject	16	Do not reject	8	Do not reject	4,183
(xi)	Management score	Reject	14	Reject	7	Reject	2,774
(xii)	Innovation score	Do not reject	11	Reject	6	Reject	1,409
(xiii)	Reported - book sales	Do not reject	16	Do not reject	8	Do not reject	4,152
(xiv)	Tax/Sales	Do not reject	16	Do not reject	8	Do not reject	4,178

Note: Each row (i)–(xiv) reports results from a separate regression corresponding to specifications (1)–(14) in Table 7 of Cai and Szeidl (2017). The attribute variable $EmpSize_i$ is measured as the log baseline number of employees for each firm. The first column reports the specification number, y reports the dependent variable. The "Original result" column reports the conclusions reached in the original paper for the corresponding null hypothesis of no peer effects. The "$\boldsymbol{c}$ test" examines dependence operating through peers' attributes, while the "$\boldsymbol{cy}$ test" augments the instrument set with functions of peers' outcomes. The columns p_c and p_{cy} report the number of basis functions used in each test, n reports the sample size. Peer exposure is measured using a simple average social vector, $w_{i,t}$, where each player is exposed to the average attribute/outcome of peers within the same meeting group at time t. All specification controls are identical to those reported in Table E2 in Appendix E.2. Tests are performed at the 95% confidence level. Standard errors are clustered at the meeting group level.

but also through outcome-based spillovers. The combined new evidence points to spillovers particularly for performance measures and intermediate mechanisms linked to information sharing and managerial practices.

4.3 Student achievement (Booij et al., 2017)

4.3.1 Context and findings

In our third demonstration we revisit the paper by Booij et al. (2017), who study a large-scale field experiment conducted at the University of Amsterdam in the Netherlands. The intervention randomly assigned first-year economics students to tutorial groups. By exogenously varying group composition, this design created substantial variation to estimate how students' performance responds to changes in the academic environment.[13]

The authors study how students' performance (measured by the number of collected credits) is affected by prior achievement of the peers in the same tutorial group (measured by their secondary-school GPA). Their main results are reproduced in Appendix E, Table E3. From columns (1) to (5), academic performance is regressed on various summary statistics of peers' GPA (e.g., the mean, dispersion, and their interaction), using a sequence of increasingly rich specifications that allow for heterogeneous responses across the ability distribution. As shown in Table E3, these linear regressions find limited evidence that peer achievement or peer heterogeneity affects student outcomes, except for some specifications involving higher-order interactions.[14]

4.3.2 Test results

In Table 3 we apply our nonparametric test to re-examine the peer effects documented in Booij et al. (2017). Each row of Table 3 corresponds, in order of appearance, to a specification in Table E3. Since peer exposure variables are specification-specific, we report them as $w'c$ and $w'y$ respectively. The remaining columns (e.g. original results, $\boldsymbol{c}$ and $\boldsymbol{cy}$ tests, p_c and p_{cy}, the number of peer attribute terms l) follow the conventions of the previous tables.

[13] Students attended weekly tutorials with the same group throughout the course, and all teaching assistants followed a common syllabus, minimizing the chance that differences in outcomes could be attributed to instruction quality.

[14] As clarified in Appendix E "higher-order" interactions refer to interaction terms involving multiple variables (e.g. peer mean × peer dispersion × own GPA).

Table 3: Peer Exposure Design and Test Results: Booij et al. (2017)

Dependent variable: Number of credits achieved

n =1,876

	Peer exposure in attributes					Peer exposure in outcomes		
	Peer exposure $w'c$	l	Original result	p_c	$\boldsymbol{c}$ test	Peer exposure $w'y$	p_{cy}	$\boldsymbol{cy}$ test
(i)	$w'_{i,\text{avg}}GPA$	1	Do not reject	12	Do not reject	$w'_{i,\text{avg}}y$	6	Reject
(ii)	$w'_{i,\text{avg}}GPA$	1	Do not reject	12	Reject	$w'_{i,\text{avg}}y$	6	Reject
		l, ℓ					$p_{cy,m}$	$\boldsymbol{cy}$ test
(iii)	$w'_{i,\text{avg}}GPA$, $w'_{i,\text{sd}}GPA$	2, 2	Do not reject	6	Reject	$w'_{i,\text{avg}}y$, $w'_{i,\text{sd}}y$	3	Reject
(iv)	$w'_{i,\text{avg}}GPA$, $w'_{i,\text{sd}}GPA$, $w'_{i,\text{avg}}GPA \times w'_{i,\text{sd}}GPA$	3, 3	Reject	4	Reject	$w'_{i,\text{avg}}y$, $w'_{i,\text{sd}}y$, $w'_{i,\text{avg}}y \times w'_{i,\text{sd}}y$	2	Reject
(v)	$w'_{i,\text{avg}}GPA$, $w'_{i,\text{sd}}GPA$, $w'_{i,\text{avg}}GPA \times w'_{i,\text{sd}}GPA$, $GPA_i \times w'_{i,\text{avg}}GPA$, $GPA_i \times w'_{i,\text{sd}}GPA$, $GPA_i \times w'_{i,\text{avg}}GPA \times w'_{i,\text{sd}}GPA$	6, 2	Reject	2	Reject	$w'_{i,\text{avg}}y$, $w'_{i,\text{sd}}y$	2	Reject

Note: Each row corresponds to specifications (i)–(v) in Table 4 of Booij et al. (2017). Columns 1-3 report the specification number, the construction of peer attribute exposure variables, and the number of peer attribute terms (l). The "Original result" column reports the conclusions reached in the original paper for the corresponding null hypothesis of no peer effects. The "$\boldsymbol{c}$ test" examines dependence operating through peers' attributes, while the "$\boldsymbol{cy}$ test" augments the test with functions of peers' outcomes. The columns p_c and p_{cy} report the number of basis functions used in each test. The outcome variable, student performance, is measured by the number of credits obtained in the first year of university. The ability attribute, GPA, refers to the vector of students' pre-university GPA. In specifications (i) and (ii), peer attribute exposure is measured using a weighted-average vector, $w_{i,\text{avg}}$, where each student is exposed to the average GPA of peers within the same tutorial group. In specifications (iii) and (iv), peer exposure additionally incorporates dispersion in peer characteristics through a standard-deviation vector, $w_{i,\text{sd}}$, which captures variation in peers' GPA within the group with specification (iv) additionally including their interaction. Accordingly the third column now reports l, ℓ and the second last column the recommended $p_{cy,m}$ from (B.5). In specification (v), the peer exposure terms in (iv) are further reweighted by individual GPA, interacting peer exposure with the student's own GPA to allow the strength of spillovers to vary with academic ability. All specification controls are identical to those reported in Table E3 in Appendix E.3. Tests are performed at the 95% confidence level. Standard errors are clustered at the tutorial-group level.

The comparison with the results by Booij et al. (2017) reveals a divergence between our nonparametric test and the authors' original findings. Our $\boldsymbol{c}$ test rejects the null in specifications (ii) and (iii), whereas the original regressions detect no significant linear peer effects. The other results, for specifications (i), (iv) and (v) are consistent with the authors' conclusion. Our $\boldsymbol{cy}$ test additionally detects spillovers in specification (i), while the original results of the authors detected no significant spillovers. Overall, our tests suggest that dependence arising from peer academic composition may operate in nonlinear ways that a linear model does not fully capture, but are approximated by polynomial terms.

5 Conclusion

Cross-unit dependence has long been recognized as a central concern in economics, reflecting both fundamental identification problems and first-order implications for econometric inference (Manski, 1993; Conley, 1999). This paper proposes a novel nonparametric test for spillovers operating through peers' attributes and/or outcomes, and provides a full asymptotic theory for it. The test has several appealing features. First, it only requires estimation under the null hypothesis of no spillovers, thereby avoiding nonparametric estimation altogether. Second, it is versatile, accommodating a wide range of data structures, including settings in which the interaction structure is incomplete or measured with error.

Our approach complements existing methods by offering a simple diagnostic to assess whether cross-unit dependence is present and whether linear approximations are likely to be informative. We illustrate its usefulness through three empirical applications, which suggest that the test can uncover forms of cross-unit dependence that are missed by standard specifications. More broadly, our results emphasize that modeling choices regarding spillovers and interference can be guided by empirical evidence, rather than imposed a priori through restrictive functional or informational assumptions.

References

Acemoglu, D., A. E. Özdaglar, and A. Tahbaz-Salehi (2016). Networks, shocks, and systemic risk. In Y. Bramoullé, A. Galeotti, and B. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 569–610. Oxford University Press.

Arduini, T., E. Patacchini, and E. Rainone (2020). Treatment effects with heterogeneous externalities. *Journal of Business & Economic Statistics 38*, 826–838.

Avvisati, F., M. Gurgand, N. Guyon, and E. Maurin (2013). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies 81*, 57–83.

Baird, S., J. A. Bohren, C. McIntosh, and B. Özler (2018). Optimal design of experiments in the presence of interference. *Review of Economics and Statistics 100*, 844–860.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*, 1047–1094.

Bandiera, O. and I. Rasul (2006). Social networks and technology adoption in northern Mozambique. *Economic Journal 116*, 869–902.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*, 1236498.

Bobonis, G. J. and F. Finan (2009). Neighborhood peer effects in secondary school enrollment decisions. *Review of Economics and Statistics 91*, 695–716.

Booij, A. S., E. Leuven, and H. Oosterbeek (2017). Ability peer effects in university: Evidence from a randomized experiment. *Review of Economic Studies 84*, 547–579.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*, 41–55.

Breza, E., A. G. Chandrasekhar, T. H. McCormick, and M. Pan (2020). Using aggregated relational data to feasibly identify network structure without network data. *American Economic Review 110*, 2454–84.

Brown, K. M. and R. A. Laschever (2012). When they're sixty-four: Peer effects and the timing of retirement. *American Economic Journal: Applied Economics 4*, 90–115.

Bursztyn, L. and S. Fiorin (2017). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *American Economic Review 107*, 2412–2443.

Cai, J., A. De Janvry, and E. Sadoulet (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics 7*, 81–108.

Cai, J. and A. Szeidl (2017). Interfirm relationships and business performance. *Quarterly Journal of Economics 133*, 1229–1282.

Chen, X. (2007). *Large sample sieve estimation of semi-nonparametric models, In: Handbook of Econometrics*, Volume 6B, Chapter 76, pp. 5549–5632. North Holland.

Comola, M., R. Dieye, and B. Fortin (2025). Heterogeneous peer effects and gender-based interventions for teenage obesity. *Journal of Health Economics 102*, 103023.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*, 1–45.

Duflo, E. and E. Saez (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics 118*, 815–842.

Gupta, A. (2018). Nonparametric specification testing via the trinity of tests. *Journal of Econometrics 203*, 169–185.

Gupta, A., J. Lee, and F. Rossi (2025). Testing linearity of spatial interaction functions à la Ramsey. arXiv:2412.14778 [econ.EM].

Gupta, A., X. Qu, S. Srisuma, and J. Zhang (2025). Wald inference on varying coefficients. arXiv:2502.03084 [econ.EM].

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics 1*, 34–68.

Hirano, K. and J. Hahn (2010). Design of randomized experiments to measure social interaction effects. *Economics Letters 106*, 51–53.

Hong, Y. and H. White (1995). Consistent specification testing via nonparametric series regression. *Econometrica 63*, 1133–1159.

Hsieh, C.-S. and X. Lin (2017). Gender and racial peer effects with endogenous network formation. *Regional Science and Urban Economics 67*, 135–147.

Hudgens, M. G. and M. E. Halloran (2008). Toward causal inference with interference. *Journal of the American Statistical Association 103*, 832–842.

Jenish, N. (2012). Nonparametric spatial regression under near-epoch dependence. *Journal of Econometrics 167*, 224–239.

Jenish, N. (2016). Spatial semiparametric model with endogenous regressors. *Econometric Theory 32*, 714–739.

Jenish, N. and I. R. Prucha (2012). On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics 170*, 178 – 190.

Jochmans, K. (2023). Peer effects and endogenous social interactions. *Journal of Econometrics 235*, 1203–1214.

Lalive, R. and M. A. Cattaneo (2009). Social interactions and schooling decisions. *Review of Economics and Statistics 91*, 457–477.

Lee, J. and P. M. Robinson (2016). Series estimation under cross-sectional dependence. *Journal of Econometrics 190*, 1–17.

Liu, L. and M. G. Hudgens (2014). Large sample randomization inference of causal effects in the presence of interference. *Journal of the American Statistical Association 109*, 288–301.

Lubold, S., A. G. Chandrasekhar, and T. H. McCormick (2023). Identifying the latent space geometry of network models through analysis of curvature. *Journal of the Royal Statistical Society Series B: Statistical Methodology 85*, 240–292.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*, 531–542.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*, 112–45.

McNealis, V., E. E. M. Moodie, and N. Dean (2024). Revisiting the effects of maternal education on adolescents' academic performance: Doubly robust estimation in a network-based observational study. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 73*, 715–734.

Miguel, E. and M. Kremer (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*, 159–217.

Oster, E. and R. Thornton (2012). Determinants of technology adoption: Peer effects in menstrual cup take-up. *Journal of the European Economic Association 10*, 1263–1293.

Ozier, O. (2018). Exploiting externalities to estimate the long-term effects of early childhood deworming. *American Economic Journal: Applied Economics 10*, 235–62.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association 102*, 191–200.

Sacerdote, B. (2001). Peer effects with random assignment: Results from Dartmouth roommates. *Quarterly Journal of Economics 116*, 681–704.

Vazquez-Bare, G. (2023). Identification and estimation of spillover effects in randomized experiments. *Journal of Econometrics 237*, 105237.

Viviano, D., L. Lei, G. Imbens, B. Karrer, O. Schrijvers, and L. Shi (2025). Causal clustering: Design of cluster experiments under network interference. arXiv:2502.03084 [econ.EM].

Xu, X. and L.-f. Lee (2015). A spatial autoregressive model with a nonlinear transformation of the dependent variable. *Journal of Econometrics 186*, 1–18.

Xu, X. and L.-f. Lee (2018). Sieve maximum likelihood estimation of the spatial autoregressive Tobit model. *Journal of Econometrics 203*, 96–112.

Young, H. P. (2009). Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review 99*, 1899–1924.

A Simulation study

This section reports Monte Carlo results on the empirical size of our test. We set $k = 3$ and $\beta = (0.5, -2, 1)'$. The regressor matrix X is $n \times k$ with ones in the first column and, in each replication, the remaining columns drawn independently across units from $U[-2, 2]$ and $U[-2.5, 2.5]$. Outcomes are generated by $y_i = x_i'\beta + \epsilon_i$, with no outcome or covariate spillovers. We implement the ***cy*** test using Hermite polynomials as the basis functions $\psi_j(\cdot)$ and (3.8) as the selection rule for p_{cy}, with $l = 2$.

For the cluster-correlation specification paired with cluster-robust standard errors, stack $\epsilon = (\epsilon'_{(1)}, \ldots, \epsilon'_{(G)})'$, where $\epsilon_{(g)} \in \mathbb{R}^{n_g}$. Then

$$var(\epsilon) = \Sigma = diag\left[\Sigma_1, \ldots, \Sigma_G\right], \qquad \Sigma_g = \frac{1}{2} I_{n_g} + \frac{1}{2} \mathbf{1}_{n_g} \mathbf{1}'_{n_g}, \tag{A.1}$$

where $\mathbf{1}_{n_g}$ is the $n_g \times 1$ vector of ones. Hence $var(\epsilon_i) = 1$, $cov(\epsilon_i, \epsilon_j) = 1/2$ for distinct units in the same cluster, and $cov(\epsilon_i, \epsilon_j) = 0$ across clusters. Draws satisfy $\epsilon = L\zeta$ with $LL' = \Sigma$ and $\zeta \in \mathbb{R}^n$ having independent $N(0, 1)$ or t_{10} components in the same cluster. Clusters partition indices into contiguous blocks: the first $G - 1$ blocks have size $\lfloor n/G \rfloor$, and the Gth block contains all remaining units.

We report empirical rejection frequencies under H_0 at nominal level 5%, with 1000 replications, for $n = 100, 200, 400, 700, 1000$ in non-lattice designs and $n = 100, 210, 400, 702, 992$ on a two-dimensional lattice (see below for details). We report results with χ^2_q critical values as well as the asymptotic one based on $N(0, 1)$. Gupta et al. (2025) observe that the former can control size better in smaller samples. Following Gupta et al. (2025), we consider five designs for W, all normalized by spectral norm.

(1) *Exponential distance.*

$$w_{ij} = \exp\left(-|s_i - s_j|\right) \mathbb{1}\{|s_i - s_j| < \log n\},$$

where s_i is the location of unit i on $[0, n]$, and $s_i \sim$i.i.d. $U[0, n]$.

(2) *Cutoff.*

$$w_{ij} = \Phi(-d_{ij}) \, \mathbb{1}\{c_{ij} < n^{-2/3}\},$$

with Φ the standard normal c.d.f., $d_{ij} \sim U[-3, 3]$, $c_{ij} \sim U[0, 1]$ i.i.d..

(3) *Circulant.* $W_{i,i-1} = W_{i,i+1} = 1/2$, $i = 1, \ldots, n$.

(4) *Random.* W is symmetric with entries in $\{0, 1\}$ and a total of $\lfloor 2n^{6/5} \rfloor$ non-zero elements off the diagonal. The average number of neighbors per row lies between 5.9 and 9.2 over the values of n we use.

(5) *Lattice.* Take (m_1, m_2) with $n = m_1 m_2$, e.g. $(10, 10), (14, 15), (20, 20), (26, 27), (31, 32)$, yielding $n = 100, 210, 400, 702$ and 992. Then generate

$$w_{m_2(k-1)+j,\, m_2(k-1)+j+1} = w_{m_2(k-1)+j,\, m_2(k-1)+j-1} = 1$$

for admissible (k, j), and all other entries equal to zero.

Table A1: Monte Carlo size

			χ^2_q critical values					$N(0,1)$ critical value				
ζ	n	p	expo	cutoff	circ	rand	latt	expo	cutoff	circ	rand	latt
$N(0,1)$	100	2	0.036	0.037	0.053	0.044	0.052	0.058	0.061	0.085	0.066	0.074
	200	3	0.047	0.037	0.084	0.042	0.050	0.071	0.059	0.109	0.065	0.081
	400	4	0.031	0.024	0.062	0.030	0.046	0.051	0.037	0.089	0.047	0.066
	700	4	0.039	0.037	0.059	0.047	0.047	0.057	0.053	0.084	0.072	0.066
	1000	5	0.030	0.038	0.052	0.022	0.049	0.045	0.052	0.076	0.042	0.071
t_{10}	100	2	0.040	0.049	0.049	0.042	0.046	0.062	0.078	0.081	0.070	0.075
	200	3	0.040	0.032	0.070	0.039	0.054	0.062	0.057	0.099	0.072	0.087
	400	4	0.038	0.034	0.064	0.040	0.045	0.060	0.058	0.097	0.055	0.067
	700	4	0.037	0.049	0.064	0.032	0.052	0.052	0.062	0.088	0.051	0.074
	1000	5	0.021	0.044	0.068	0.028	0.043	0.044	0.056	0.090	0.051	0.066

Note: Critical values are based on χ^2_q distributions and the standard normal distribution. Nominal size is 5%. All results are computed using cluster–robust variance. The data generating process is simulated under the null hypothesis of no spillovers. The reported rejection frequencies correspond to the ***cy*** test. For the lattice design, the sample sizes are $n = 100, 210, 400, 702$ and 992.

The results are reported in Table A1. For disturbances generated via $N(0, 1)$ shocks, there appears to be some benefit from using the χ^2_q critical values even at $(n, p) = (1000, 5)$ for the circulant and lattice cases. For the t_{10} case, this remains true for the circulant case but not for the lattice case. In the remaining four designs for W, the results follow broadly the same pattern regardless of $N(0,1)$ or t_{10} shocks: χ_q^2 critical values can sometimes, though not always, control size better than the $N(0,1)$ one for small (n,p), but this effect diminishes as (n,p) grow. These results are on expected lines.

B Extensions

B.1 Heterogeneous cross-unit dependence

It is straightforward to extend our method to allow for multiple channels of cross-unit dependence, i.e. multiple social weight matrices, but we present our theory for the single channel case in (2.1) for notational simplicity.[15] Indeed, suppose that we have ℓ channels of social dependence, each encoded in a weight matrix W_h, $h = 1, \ldots, \ell$. Then we can write the model

$$y_i = \sum_{h=1}^{\ell} f_h\left(w'_{h,i}y\right) + x'_i\beta + \sum_{h=1}^{\ell}\sum_{j=1}^{l} g_{hj}\left(w'_{h,i}c_j\right) + \epsilon_i, \quad i = 1, ..., n, \tag{B.1}$$

where $f_h(\cdot)$ and $g_{hj}(\cdot)$ from $\mathbb{R}$ to $\mathbb{R}$ are $\ell(l+1)$ unknown functions and $W_h = (w_{h,1}, \ldots, w_{h,n})'$ are social weight matrices that are either fixed or exogenous and have zero diagonals, $h = 1, \ldots, \ell$ and $j = 1, \ldots, l$. The null of interest is then

$$\mathcal{H}_0 : f_h(s) = 0 \text{ and } g_{hj}(s) = 0,\ h = 1, \ldots, \ell,\ j = 1, \ldots, l, \tag{B.2}$$

for all $s \in \textit{support}(s)$.

This can be approximated by series approximations exactly as in (2.5) and (2.6) albeit with more subscripting, as we now show. Now we have the approximations

$$f_h(s) = \sum_{i=1}^{p} \mu_{f_{h,i}}\psi_i(s) + r_{f_h}(s), g_{hj}(s) = \sum_{i=1}^{p} \mu_{g_{hj,i}}\psi_i(s) + r_{g_{hj}}(s), h = 1, \ldots, \ell, j = 1, \ldots, l, \tag{B.3}$$

with $\mu_f = (\mu_{f_1,1}, \ldots, \mu_{f_1,p}, \ldots, \mu_{f_\ell,1}, \ldots, \mu_{f_\ell,p})'$, $\mu_{g_j} = \left(\mu_{g_{1j},1}, \ldots, \mu_{g_{1j},p}, \mu_{g_{\ell j},1}, \ldots, \mu_{g_{\ell j},p}\right)'$ vectors of unknown series coefficients and $r_{f_h}(s)$, $r_{g_{hj}}(s)$ approximation errors. We define

[15] Social interaction models with multiple social weight matrices have been characterized under a range of alternative assumptions (Hsieh and Lin, 2017; Arduini et al., 2020; Comola, Dieye, and Fortin, 2025).

our approximate null hypothesis as

$$\mathcal{H}_{0A} : \mu_f = 0 \text{ and } \mu_{g_j} = 0,\ j = 1, \ldots, l,\ \text{for some } \beta. \tag{B.4}$$

Now, define the $n \times 1$ vector $\Upsilon_{f_h,i}(y) = \left(\psi_i(w'_{h,1}y), \ldots, \psi_i(w'_{h,n}y)\right)'$ and the $n \times 1$ vectors $\Upsilon_{g_{hj},i}(c_j) = \left(\psi_i(w'_{h,1}c_j), \ldots, \psi_i(w'_{h,n}c_j)\right)'$, for each $i = 1, \ldots, p$, $h = 1, \ldots, \ell$ and $j = 1, \ldots, l$. Next, concatenate these to write $\Upsilon_{f,g}$ as before. This is now an $n \times q$ matrix, where $q = p\ell(l+1)$. We can now proceed as before in the $\ell = 1$ case with these objects as the test statistic building blocks, noting that now we recommend

$$p_{cy,m} = \frac{\left[n^{1/3}\right]}{\ell + l},\ \text{both } \boldsymbol{c} \text{ and } \boldsymbol{y} \text{ spillovers.} \tag{B.5}$$

B.2 Embedded graphs

We now consider a setting in which the observed social matrix is generated by an embedded graph model rather than treated as exogenous. By embedding the observed graph into a latent social space, we interpret it as a noisy measurement of an underlying latent structure. This reframing is particularly valuable in applied work, where social interaction data are often incomplete or measured with error, yet the latent structure remains informative to the researcher. When social interactions are mis-measured, our approach interprets these errors as small perturbations relative to the structurally generated network. We then derive formal conditions under which such perturbations are asymptotically negligible, ensuring the validity of inference based on the observed data.

Let us assume the researcher observes a network encoded by the $n \times n$ matrix $W(\kappa, z)$, the elements $w_{ij}(\kappa, z_o, z_l)$ of which are indicator functions that take the value unity if the unknown parameter vector $\kappa \in \mathscr{K}$, and the observed covariate vector z_o and latent vector z_l satisfy some prescribed condition. This specification encompasses standard network formation models with link functions that permit consistent estimation of κ, such as exponential link functions. The interpretation is that the adjacency matrix $W(\kappa, z)$ represents an embedding of the graph in a latent social space, rather than a fixed object observed without error. Such formulations generalize link formation based solely on observed covariates to settings where proximity in latent space governs tie formation, see for example Breza, Chandrasekhar, McCormick, and Pan (2020) and Lubold, Chandrasekhar, and McCormick (2023).

Accordingly we assume that, conditional on the latent variable vector z_l, the researcher has access to estimates $\hat{w}_{ij} = w_{ij}\left(\hat{\kappa}, z_o\right)$, where $\hat{\kappa}$ is some estimate of the parameterization of the link function that defines the probability of $w_{ij} = 1$ as a function of κ, z_o and z_l. Furthermore, we assume that there exists a sequence $s_n = s \to \infty$ such that

$$\|\hat{\kappa} - \kappa\| = O_p\left(s^{-1}\right), \tag{B.6}$$

and that this rate of convergence carries over to the maximum row-sum of $W\left(\hat{\kappa}, z\right) - W(\kappa, z)$, i.e.

$$\sup_{j=1,\ldots,n} \sum_{i=1}^{n} \left(\hat{w}_{ij} - w_{ij}\right) = O_p\left(s^{-1}\right). \tag{B.7}$$

This condition is motivated by the observation that adjacency matrices require some control over their norms to limit dependence to a manageable degree. Row and column summability is a typical assumption. Indeed, if $w_{ij}(\kappa, z)$ were a differentiable function in κ we could use the mean value theorem to write $\hat{w}_{ij} - w_{ij} = \left.\frac{\partial w_{ij}(\kappa,z)}{\partial \kappa}\right|_{\kappa=\bar{\kappa}}' \left(\hat{\kappa} - \kappa\right)$ for an intermediate point $\bar{\kappa}$ and obtain (B.7) if we assume row-summability of the derivative matrix, i.e.

$$\sup_{j=1,\ldots,n} \sup_{\kappa \in \mathscr{K}} \sum_{i=1}^{n} \frac{\partial w_{ij}(\kappa, z)}{\partial \kappa} = O_p(1),$$

uniformly in z. Of course in our case $\hat{w}_{ij}$ and w_{ij} are non-differentiable because they are indicator functions, hence the condition (B.7). For example, if each unit only has a fixed number of neighbours, as in a 'k nearest neighbours' setup, then (B.7) will be satisfied as long as

$$\hat{w}_{ij} - w_{ij} = O_p\left(s^{-1}\right), \tag{B.8}$$

because the sum on the LHS of (B.7) will have only a fixed number of non-zero summands.

Now, for each $i = 1, \ldots, p$, define the $n \times 1$ vector $\hat{\Upsilon}_{f,i}(y) = (\psi_i(\hat{w}_1' y), \ldots, \psi_i(\hat{w}_n' y))'$ and the $n \times 1$ vectors $\hat{\Upsilon}_{g_j,i}(c_j) = (\psi_i(\hat{w}_1' c_j), \ldots, \psi_i(\hat{w}_n' c_j))'$, where $\hat{w}_i$ has elements $\hat{w}_{ij}$, $j = 1, \ldots, n$, and write

$$\hat{\Upsilon}_{f,g} = \begin{pmatrix} \hat{\Upsilon}_{f,1}(y) & \ldots & \hat{\Upsilon}_{f,p}(y) & \hat{\Upsilon}_{g_1,1}(c_1) & \ldots & \hat{\Upsilon}_{g_1,p}(c_1) & \ldots & \hat{\Upsilon}_{g_l,1}(c_l) & \ldots & \hat{\Upsilon}_{g_l,p}(c_l) \end{pmatrix},$$

which is an $n \times q$ matrix, where $q \sim p$ asymptotically. Then, we can apply our test as long as the estimates $\hat{w}_{ij}$ satisfy

$$\left\| n^{-1/2}\hat{\Upsilon}_{f,g} - n^{-1/2}\Upsilon_{f,g} \right\| = o_p(1), \tag{B.9}$$

conditional on z_l. The squared LHS of (B.9) is bounded by a sum of nq terms of the type

$$n^{-1}\left(\psi\left(\hat{w}'\ell\right) - \psi\left(w'\ell\right)\right)^2, \tag{B.10}$$

where we omit subscripting for brevity and let ℓ denote a generic observed $n \times 1$ vector with components ℓ_i such that $\ell_i = O_p(1)$. Assuming that the generic basis function $\psi(\cdot)$ is differentiable with derivative $\psi'(\cdot)$ and $E(\psi'(x))^2 < C$, with C a generic constant, we use the mean value theorem and (B.7) to observe that (B.10) is

$$\begin{aligned} n^{-1}\psi'(\bar{x})^2\left((\hat{w}-w)'\,\ell\right)^2 &= O_p\left(n^{-1}\right)\cdot\left(\sum_{i=1}^{n}\left(\hat{w}_i - w_i\right)\ell_i\right)^2 \\ &= O_p\left(n^{-1}\right)\cdot\left(\sum_{i=1}^{n}\left(\hat{w}_i - w_i\right)\right)^2 \\ &= O_p\left(s^{-2}n^{-1}\right), \end{aligned} \tag{B.11}$$

where $\hat{w}'\ell \leq \bar{x} \leq w'\ell$ and $\hat{w}_i$ and w_i are elements of the $n \times 1$ vectors $\hat{w}$ and w, respectively. Then we conclude that

$$\left\| n^{-1/2}\hat{\Upsilon}_{f,g} - n^{-1/2}\Upsilon_{f,g} \right\|^2 = O_p\left(ps^{-2}\right), \tag{B.12}$$

so that (B.9) holds if $p^{1/2}s^{-1} \to 0$.

C Proofs of Theorems

C.1 Preliminary results

Theorem C1. *Under Assumptions 1-4, under $\mathcal{H}_{0A}$ in (2.6), for $p^3/n \to 0$ as $n \to \infty$,*

$$\left\| \hat{d} - d \right\| = O_p\left(\frac{p^{3/2}}{n}\right). \tag{C.1}$$

Proof. We first establish a preliminary bound. Let $1_g(i,j)$ be and indicator function that takes the value 1 when i and j are in the same cluster g and zero otherwise. Observe that, by Assumptions 1, 3 and $m \sim p$,

$$\begin{aligned}
\mathbb{E}\left\|n^{-1}Z'\epsilon\right\|^2 &= n^{-2}\sum_{i=1}^{n}\sigma_i^2\mathbb{E}\left\|z_i\right\|^2 + n^{-2}\sum_{g=1}^{G}\sum_{i\neq j}1_g(i,j)\sigma_{ij}\mathbb{E}z_{ig}'z_{jg} \\
&\leq Kn^{-1}m + n^{-2}\sum_{g=1}^{G}\sum_{i\neq j}1_g(i,j)\sigma_{ij}\left(\mathbb{E}\left\|z_{ig}\right\|^2\right)^{1/2}\left(\mathbb{E}\left\|z_{jg}\right\|^2\right)^{1/2} \\
&= O\left(p\left(n^{-1}+n^{-2}\sum_{g=1}^{G}n_g^2\right)\right) = O\left(pn^{-1}\right), \qquad \text{(C.2)}
\end{aligned}$$

where the last equality follows because $n^{-2}\sum_{g=1}^{G}n_g^2 = O\left(n^{-2}G\right) = O\left(n^{-1}\right)$, recalling that $\sup_{g=1,\ldots,G}n_g < K$ and so G and n have the same asymptotic order. Thus, by the Markov inequality,

$$\left\|n^{-1}Z'\epsilon\right\| = O_p\left(\sqrt{\frac{p}{n}}\right). \qquad \text{(C.3)}$$

We also note that, under Assumptions 2, 3 and 4,

$$\left\|\hat{\beta}-\beta\right\| = \left\|\left(\frac{1}{n}X'\mathcal{P}_ZX\right)^{-1}\frac{1}{n}X'\mathcal{P}_Z\epsilon\right\| = O_p\left(\left\|\frac{Z'\epsilon}{n}\right\|\right). \qquad \text{(C.4)}$$

Let $R = (r(w_1'y),\ldots,r(w_n'y))'$ be the $n\times 1$ vector of approximation errors in (2.5) with $R_i = r(w_i'y)$. From the 2SLS expression for $\hat{\beta}-\beta$ in (C.4),

$$\begin{aligned}
\hat{d} &= -\frac{2}{n}U'\mathcal{P}_Z\left(I - X(X'\mathcal{P}_ZX)^{-1}X'\mathcal{P}_Z\right)\epsilon - \frac{2}{n}U'\mathcal{P}_ZR \\
&= -\frac{2}{n}U'\mathcal{P}_Z\left(I - \mathcal{P}_ZX(X'\mathcal{P}_ZX)^{-1}X'\mathcal{P}_Z\right)\mathcal{P}_Z\epsilon - \frac{2}{n}U'\mathcal{P}_ZR \\
&= -\frac{2}{n}\hat{J}'\hat{M}^{-1/2}\left(I - \hat{M}^{-1/2}\hat{N}\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\hat{N}'\hat{M}^{-1/2}\right)\hat{M}^{-1/2}Z'\epsilon - \frac{2}{n}\hat{J}'\hat{M}^{-1}Z'R \\
&= -\frac{2}{n}\hat{J}'\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}Z'\epsilon - \frac{2}{n}\hat{J}'\hat{M}^{-1}Z'R, \qquad \text{(C.5)}
\end{aligned}$$

where $\hat{\mathcal{K}}_{NM} = \left(I - \hat{M}^{-1/2}\hat{N}\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\hat{N}'\hat{M}^{-1/2}\right)$.

From (2.9), we write

$$\left\|\hat{d}-d\right\| \leq \left\|\frac{2}{n}\hat{J}'\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}Z'\epsilon - \frac{2}{n}J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}Z'\epsilon\right\| + \left\|\frac{2}{n}\hat{J}'\hat{M}^{-1}Z'R\right\|. \tag{C.6}$$

Subsequently, denote $\Delta_B^A = A - B$ for conformable A and B. Then, via some standard albeit tedious algebra, the first term on the RHS of (C.6) is bounded by

$$\begin{aligned}
&\left\|\Delta_J^{\hat{J}}\right\|\left\|\hat{M}^{-1}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| + \|J\|\left\|\Delta_{M^{-1}}^{\hat{M}^{-1}}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| \\
&+\left\|\Delta_J^{\hat{J}}\right\|\left\|\hat{M}^{-1}\right\|\left\|\hat{N}\right\|\left\|\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\right\|\left\|\hat{N}\right\|\left\|\hat{M}^{-1}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| \\
&+\|J\|\left\|\Delta_{M^{-1}}^{\hat{M}^{-1}}\right\|\left\|\hat{N}\right\|\left\|\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\right\|\left\|\hat{N}\right\|\left\|\hat{M}^{-1}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| \\
&+\|J\|\left\|M^{-1}\right\|\left\|\Delta_N^{\hat{N}}\right\|\left\|\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\right\|\left\|\hat{N}\right\|\left\|\hat{M}^{-1}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| \\
&+\|J\|\left\|M^{-1}\right\|\|N\|\left\|\Delta_{(N'M^{-1}N)^{-1}}^{(\hat{N}'\hat{M}^{-1}\hat{N})^{-1}}\right\|\left\|\hat{N}\right\|\left\|\hat{M}^{-1}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| \\
&+\|J\|\left\|M^{-1}\right\|\|N\|\left\|\left(N'M^{-1}N\right)^{-1}\right\|\left\|\Delta_N^{\hat{N}}\right\|\left\|\hat{M}^{-1}\right\|\left\|\frac{1}{n}Z'\epsilon\right\| \\
&+\|J\|\left\|M^{-1}\right\|\|N\|\left\|\left(N'M^{-1}N\right)^{-1}\right\|\|N\|\left\|\Delta_{M^{-1}}^{\hat{M}^{-1}}\right\|\left\|\frac{1}{n}Z'\epsilon\right\|
\end{aligned} \tag{C.7}$$

Under Assumption 4, we have

$$\left\|\Delta_N^{\hat{N}}\right\| = O_p\left(\frac{p}{\sqrt{n}}\right) \quad \text{and} \quad \left\|\Delta_J^{\hat{J}}\right\| = O_p\left(\frac{p}{\sqrt{n}}\right). \tag{C.8}$$

Also, under Assumptions 3 and 4,

$$\left\|\Delta_{M^{-1}}^{\hat{M}^{-1}}\right\| \leq \left\|M^{-1}\right\|\left\|\hat{M}^{-1}\right\|\left\|\Delta_M^{\hat{M}}\right\| = O_p\left(\frac{p}{\sqrt{n}}\right) \tag{C.9}$$

and similarly, under Assumptions 3 and 4,

$$\left\|\Delta_{(N'M^{-1}N)^{-1}}^{(\hat{N}'\hat{M}^{-1}\hat{N})^{-1}}\right\| \leq \left\|N'M^{-1}N'\right\|\left\|\hat{N}'\hat{M}^{-1}\hat{N}'\right\|\left\|\Delta_{N'M^{-1}N}^{\hat{N}'\hat{M}^{-1}\hat{N}}\right\|$$

$$= O_p\left(\frac{p}{\sqrt{n}}\right). \tag{C.10}$$

Thus, upon recalling (C.3), the first term at the RHS of (C.6) is observed to be $O_p\left(n^{-1}p^{3/2}\right)$. The second term at the RHS of (C.6) is instead

$$\left\| \frac{2}{n} \hat{J}' \hat{M}^{-1} Z' R \right\| = O_p\left(\frac{1}{n}\|R\|\|Z\|\right) = O_p(p^{-\nu}), \tag{C.11}$$

where the first equality at the RHS of (C.11) follows under Assumptions 3 and 4. The second equality follows since $\|Z\| = O_p(\sqrt{n})$ under Assumptions 3 and 4, and each component of the $n \times 1$ vector R is $O_p(p^{-\nu})$ by Assumption 3, and hence $||R|| = O_p(\sqrt{n}p^{-\nu})$. The last equality in (C.11) follows from Assumption 3.

Under Assumption 4, the first term in (C.6) dominates the second one as long as ν satisfies $n/p^{\nu+3/2} = o(1)$ as $n \to \infty$, which holds under Assumption 3. □

Theorem C2. *Under Assumptions 1-4, $\nu > 5/2$, under $\mathcal{H}_{0A}$ in (2.6) and $p^3/n = o(1)$,*

$$\mathcal{S} - \frac{nd'H^{-1}d - q}{\sqrt{2q}} = o_p(1), \text{ as } n \to \infty. \tag{C.12}$$

Proof. We can equivalently prove

$$\hat{d}'\hat{H}^{-1}\hat{d} - d'H^{-1}d = o_p\left(\frac{\sqrt{p}}{n}\right). \tag{C.13}$$

Write the LHS of (C.13) as

$$\left(\hat{d} - d\right)' \hat{H}^{-1}\hat{d} + d'H^{-1}(\hat{d} - d) + d'\hat{H}^{-1}\left(H - \hat{H}\right) H^{-1}\hat{d}, \tag{C.14}$$

which has norm bounded by

$$K\left\|\Delta_d^{\hat{d}}\right\|\left\|\hat{H}^{-1}\right\|\left\|\hat{d}\right\| + K\left\|\Delta_d^{\hat{d}}\right\|\left\|H^{-1}\right\|\|d\| + K\|d\|\left\|\hat{H}^{-1}\right\|\left\|\Delta_H^{\hat{H}}\right\|\left\|H^{-1}\right\|\left\|\hat{d}\right\|. \tag{C.15}$$

From Theorem C1, $\left\|\Delta_d^{\hat{d}}\right\| = O_p\left(n^{-1}p^{3/2}\right)$. Under Assumptions 3 and 4, and from (C.2) we have $\|d\| = O_p\left(\sqrt{p/n}\right)$. Also, from Theorem C1,

$$\left\|\hat{d}\right\| \le \left\|\Delta_d^{\hat{d}}\right\| + \|d\| = O_p\left(\sqrt{\frac{p}{n}}\right), \tag{C.16}$$

where the last equality is due to $p^2/n = o(1)$. Also, under Assumptions 1, 3 and 4, $\left\|\hat{H}^{-1}\right\| = O_p(1)$ and $\left\|H^{-1}\right\| = O_p(1)$. Thus, the first and second terms in (C.15) are $O_p\left(p^2/n^{3/2}\right)$, and these are $o_p\left(p^{1/2}/n\right)$ if $p^3/n = o(1)$.

Using 2SLS estimates for β_0 and proceeding as in (C.5), we can write

$$\begin{aligned}\hat{H} =& 4\hat{J}'\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}\tilde{\Phi}\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}\hat{J} + \frac{4}{n}\hat{J}'\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}Z'\epsilon R'Z\hat{M}^{-1}\hat{J} \\ &+ \frac{4}{n}\hat{J}'\hat{M}^{-1}Z'RR'Z\hat{M}^{-1}\hat{J}\end{aligned} \tag{C.17}$$

where $\tilde{\Phi} = Z'\tilde{\Sigma}Z/n$, with $\tilde{\Sigma}$ being an $n \times n$ block-diagonal matrix such that its g-th block $\tilde{\Sigma}_g$ has elements $\tilde{\Sigma}_{gij} = \epsilon_{ig}\epsilon_{jg}$. From (3.6), we write

$$\begin{aligned}\left\|\Delta_H^{\hat{H}}\right\| \le & \Big\|\hat{J}'\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}\tilde{\Phi}\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}\hat{J} \\ & - J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}\Phi M^{-1/2}\mathcal{K}_{NM}M^{-1/2}J\Big\| \\ & + \left\|\frac{4}{n}\hat{J}'\hat{M}^{-1/2}\hat{\mathcal{K}}_{NM}\hat{M}^{-1/2}Z'\epsilon R'Z\hat{M}^{-1}\hat{J}\right\| + \left\|\frac{4}{n}\hat{J}'\hat{M}^{-1}Z'RR'Z\hat{M}^{-1}\hat{J}\right\|.\end{aligned} \tag{C.18}$$

By standard algebra, the first term in (C.18) is bounded by

$$\begin{aligned}&\left\|\hat{J}'\hat{M}^{-1}\tilde{\Phi}\hat{M}^{-1}\hat{J} - J'M^{-1}\Phi M^{-1}J\right\| \\ &+ \left\|\hat{J}'\hat{M}^{-1}\hat{N}\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\hat{N}'\hat{M}^{-1}\tilde{\Phi}\hat{M}^{-1}\hat{J} - J'M^{-1}N\left(N'M^{-1}N\right)^{-1}N'M^{-1}\Phi M^{-1}J\right\| \\ &+ \Big\|\hat{J}'\hat{M}^{-1}\hat{N}\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\hat{N}'\hat{M}^{-1}\tilde{\Phi}\hat{M}^{-1}\hat{N}\left(\hat{N}'\hat{M}^{-1}\hat{N}\right)^{-1}\hat{N}\hat{M}^{-1}\hat{J} \\ &- J'M^{-1}N\left(N'M^{-1}N\right)^{-1}N'M^{-1}\Phi M^{-1}N\left(N'M^{-1}N\right)^{-1}NM^{-1}J\Big\|.\end{aligned}$$

We provide details for the first term in the last displayed expression, the others following similarly. Specifically,

$$\left\|\hat{J}'\hat{M}^{-1}\tilde{\Phi}\hat{M}^{-1}\hat{J} - J'M^{-1}\Phi M^{-1}J\right\| \le \left\|\Delta_J^{\hat{J}}\right\|\left\|\hat{M}^{-1}\right\|^2\left\|\tilde{\Phi}\right\|\left\|\hat{J}\right\|$$

$$+ \|J\| \left\| \Delta_{M^{-1}}^{\hat{M}^{-1}} \right\| \left\| \tilde{\Phi} \right\| \left\| \hat{M}^{-1} \right\| \left\| \hat{J} \right\| + \|J\| \left\| M^{-1} \right\| \left\| \Delta_{\Phi}^{\tilde{\Phi}} \right\| \left\| \hat{M}^{-1} \right\| \left\| \hat{J} \right\|$$
$$+ \|J\| \left\| M^{-1} \right\| \|\Phi\| \left\| \Delta_{M^{-1}}^{\hat{M}^{-1}} \right\| \left\| \hat{J} \right\| + \|J\| \left\| M^{-1} \right\|^2 \|\Phi\| \left\| \Delta_{J}^{\hat{J}} \right\|. \tag{C.19}$$

Under Assumptions 3 and 4, most terms can be handled as in the proof of Theorem C1, and $\left\| \Delta_{M^{-1}}^{\hat{M}^{-1}} \right\| = O_p(p/\sqrt{n})$ and $\left\| \Delta_{J}^{\hat{J}} \right\| = O_p(p/\sqrt{n})$. Focusing instead on $\left\| \Delta_{\Phi}^{\tilde{\Phi}} \right\|$, observe that

$$\left\| \Delta_{\Phi}^{\tilde{\Phi}} \right\| \leq \left\| \Delta_{\bar{\Phi}}^{\tilde{\Phi}} \right\| + \left\| \Delta_{\Phi}^{\bar{\Phi}} \right\|, \tag{C.20}$$

with $\bar{\Phi} = Z'\Sigma Z/n$. The first term in (C.20) is $\left\| Z'\Delta_{\Sigma}^{\tilde{\Sigma}} Z/n \right\|$, where the $m \times m$ matrix $Z'\Delta_{\Sigma}^{\tilde{\Sigma}} Z/n$ has typical element $n^{-1} \sum_{i,j=1}^{n} \sum_{g=1}^{G} t_{ijrsg} \left(\epsilon_{ig}\epsilon_{jg} - \sigma_{ijg} \right)$ where we write $t_{ijrsg} = 1_g(i,j) z_{irg} z_{jsg}$. This typical element has zero mean and variance

$$2n^{-2} \sum_{i,j=1}^{n} \sum_{g=1}^{G} t_{ijrsg}^2 \sigma_{ig}^2 \sigma_{jg}^2 + n^{-2} \sum_{i=1}^{n} \sum_{g=1}^{G} t_{iirsg}^2 \left(\mathbb{E}\epsilon_{ig}^4 - 3\sigma_{ig}^4 \right) = O\left(n^{-2} \sum_{g=1}^{G} n_g^2 \right) = O\left(n^{-1} \right), \tag{C.21}$$

under Assumptions 1 and 3 and, since, $m \sim p$ we therefore $\left\| \Delta_{\bar{\Phi}}^{\tilde{\Phi}} \right\| = O_p(p/\sqrt{n})$. The matrix in the norm in the second term in (C.20) has mean zero and the second moment of its squared Euclidean norm is bounded by $K\xi m^2/n^2 = O\left(p^2/n\right)$ under Assumptions 1, 3 and 4, rendering it $O_p(p/\sqrt{n})$ and thus $\left\| \Delta_{\Phi}^{\tilde{\Phi}} \right\| = O_p(p/\sqrt{n})$. We then conclude (C.19) is $O_p\left(p/\sqrt{n}\right)$. Similar steps yield that the first term on the RHS of (C.18) is $O_p\left(p/\sqrt{n}\right)$.

By similar arguments to those that led to (C.11), under Assumptions 3 and 4, the second term in (C.18) is bounded by

$$K \left\| \frac{1}{n} Z'\epsilon \right\| \|Z\| \|R\| = O_p\left(\frac{\sqrt{n}}{p^{\nu-1/2}} \right), \tag{C.22}$$

which is negligible compared to the first term in (C.18) since $n/p^{(\nu+1/2)} = o(1)$, under Assumptions 3 and 4. Similarly, the third term in (C.18) is $O_p(np^{-2\nu})$, which is negligible compared to the first term since $n^{3/2}/p^{2\nu+1} = o(1)$ as $n \to \infty$, under Assumptions 3 and 4. We conclude that

$$\left\| \Delta_{H}^{\hat{H}} \right\| = O_p\left(p/\sqrt{n}\right). \tag{C.23}$$

By Assumption 4, the last term in (C.15) is thus $O_p(p^2/n^{3/2})$, given $\|d\| = O_p(\sqrt{p/n})$ and

$\left\|\hat{d}\right\| = O_p(\sqrt{p/n})$. Hence, the second term in (C.15) is $o_p(\sqrt{p}/n)$ as long as $p^3/n = o(1)$, concluding the proof.

□

C.2 Proofs of main theorems

Proof of Theorem 1: Observe that under the clustered error dependence structure in Assumption 1, we can write $\epsilon_i = \sum_{r=1}^{n} b_{ir}\eta_r$, where η_r are i.i.d. mean zero and unit variance random variables and the $b_{ir} < K$ are finite constants that are non-zero only for the group that observation i belongs to. Then we have $\mathbb{E}\epsilon_{ig}\epsilon_{jg'} = 0$ for $g \neq g'$ and $\mathbb{E}\epsilon_{ig}\epsilon_{jg} = \sum_{r=1}^{n_g} b_{ir}b_{jr}$. Thus, we have $\sigma_{ijg} = \sum_{r=1}^{n_g} b_{ir}b_{jr}$ for $i \neq j$ and $\sigma_{ig}^2 = \sum_{r=1}^{n_g} b_{ir}^2$. Because n_g are finite and fixed, clearly only a finite number of the b_{ir} are non-zero for any given i or r. Thus we have $\sup_{r=1,\ldots,n}\sum_{i=1}^{n}|b_{ir}| + \sup_{i=1,\ldots,n}\sum_{r=1}^{n}|b_{ir}| < K$. Upon writing $\epsilon = B\eta$, where η is an $n \times 1$ vector with elements η_r and B is an $n \times n$ matrix with elements b_{ir}, we observe that

$$d = -\frac{2}{n}J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}Z'B\eta. \tag{C.24}$$

In view of Theorem C2, we know that $\mathcal{S} - \frac{nd'H^{-1}d-q}{\sqrt{2q}} = o_p(1)$ and therefore it is sufficient to show that $\frac{nd'H^{-1}d-q}{\sqrt{2q}} \xrightarrow{d} N(0,1)$, which by (C.24) boils down to showing that

$$\frac{n\eta'B'\mathscr{G}B\eta - q}{\sqrt{2q}} \xrightarrow{d} N(0,1), \tag{C.25}$$

where

$$\mathscr{G} = \frac{4}{n^2}ZM^{-1/2}\mathcal{K}_{NM}M^{-1/2}JH^{-1}J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}Z' = \frac{4}{n}Z\mathscr{A}Z',$$

say, where $\mathscr{A} = n^{-1}M^{-1/2}\mathcal{K}_{NM}M^{-1/2}JH^{-1}J'M^{-1/2}\mathcal{K}_{NM}M^{-1/2}$.

Theorem A.1 of Gupta, Qu, Srisuma, and Zhang (2025) applies if

$$\overline{eig}\mathscr{G} = O_p(1) \text{ and } \left(\underline{eig}\mathscr{G}\right)^{-1} = O_p(1), \tag{C.26}$$

and

$$g_{ij} = O_p(p/n) \text{ and } \sum_{i=1}^{n} g_{ij}^2 = O_p(p/n), \tag{C.27}$$

uniformly in i and j. The conditions in Assumption 3 ensure that (C.26) holds. To check

(C.27), observe that

$$g_{ij} = \frac{4}{n} z_i' \mathscr{A} z_j = O_p\left(\|z_i\|\|z_j\|/n\right) = O_p(p/n)$$

and

$$\sum_{j=1}^{n} g_{ij}^2 = \frac{16}{n^2} z_i' \mathscr{A} \left(\sum_{j=1}^{n} z_j z_j' \right) \mathscr{A} z_i = \frac{16}{n} z_i' \mathscr{A} \hat{M} \mathscr{A} z_i = O_p\left(\|z_i\|\|z_j\|/n\right) = O_p(p/n),$$

as desired. Then (C.25) follows by Theorem A.1 of Gupta et al. (2025). □

Proof of Theorem 2: Let $\gamma = (\mu', \beta')'$. Corresponding to $\tilde{d} = \partial \mathcal{Q}/\partial\gamma$ defined in (2.8) under $\mathcal{H}_{0A}$, we now define the unconstrained gradient vector, $(p+k)\times 1$, $\tilde{d}_U$ as

$$\tilde{d}_U(\mu, \beta, y) = -\frac{2}{n} U' \mathcal{P}_Z \left(y - X\beta\right), \tag{C.28}$$

where $\tilde{d}_U(0_{q\times 1}, \beta, y) = \tilde{d}$ defined in (2.8).

We partition ,$\hat{J} = n^{-1}U$ as $\hat{J} = (\hat{\Xi}, \hat{N})$, where $\hat{\Xi}$ and $\hat{N}$ are $m \times q$ and $m \times k$, respectively, with a similar partition for its expected value $J = (\Xi, N)$. Also, we define the $(q+k) \times (q+k)$ matrix $\hat{D} = \partial^2 \mathcal{Q}/\partial\gamma\partial\gamma'$, such that the first $q \times q$ block is given by

$$\hat{D}_{11} = \frac{2}{n} \Upsilon_{f,g}' \mathcal{P}_Z \Upsilon_{f,g} = 2\hat{\Xi}' \hat{M}^{-1} \hat{\Xi} \tag{C.29}$$

the block 1-2 (or the transposed of 2-1 block) is the $q \times k$ matrix

$$\hat{D}_{12} = \hat{D}_{21}' = \frac{2}{n} \Upsilon_{f,g}' \mathcal{P}_Z X = 2\hat{\Xi}' \hat{M}^{-1} \hat{N} \tag{C.30}$$

and the 2-2 block is the $k \times k$ matrix

$$\hat{D}_{22} = \frac{2}{n} X' \mathcal{P}_Z X = 2\hat{N}' \hat{M}^{-1} \hat{N}. \tag{C.31}$$

Under Assumption 3, $\|\hat{D}\| = O_p(1)$ and $\liminf_{n\to\infty} \underline{eig}\left(\hat{D}\right) > 0$ with inverse defined and partitioned in the usual way. Also, $\hat{D}$ does not depend on any unknowns. In line we our previous notation, we also define the corresponding limit quantities as $D_{11} = 2\Xi' M^{-1} \Xi$, $D_{12} = D_{21}' = \Xi' M^{-1} N$ and $D_{22} = 2N' M^{-1} N$.

From standard algebra, by the mean value theorem (MVT), given $\hat{d}$ in (2.9),

$$\hat{d}_p = \frac{\partial \mathcal{Q}}{\partial \mu'}\bigg|_{(0_{1\times q},\hat{\beta}')'} = \frac{\partial \mathcal{Q}}{\partial \mu'}\bigg|_{(0_{1\times q},\beta_0')'} + \hat{D}_{12}(\hat{\beta}-\beta_0)$$
$$0 = \frac{\partial \mathcal{Q}}{\partial \beta'}\bigg|_{(0_{1\times q},\hat{\beta}')'} = \frac{\partial \mathcal{Q}}{\partial \beta'}\bigg|_{(0_{1\times q},\beta_0')'} + \hat{D}_{22}(\hat{\beta}-\beta_0) \tag{C.32}$$

Thus,

$$\hat{d}_p = \left(I_q;\ -\hat{D}_{12}\hat{D}_{22}^{-1}\right)\begin{pmatrix}\frac{\partial \mathcal{Q}}{\partial \mu'}\\ \frac{\partial \mathcal{Q}}{\partial \beta'}\end{pmatrix}\Bigg|_{(0_{1\times q},\beta_0')'} = \left(I_q;\ -\hat{D}_{12}\hat{D}_{22}^{-1}\right)\tilde{d}_U(0_{q\times 1},\beta_0)$$
$$= \left(I_q;\ -\hat{D}_{12}\hat{D}_{22}^{-1}\right)\tilde{d}(\beta_0) \tag{C.33}$$

according to the definition in (C.28) and (2.8), and with I_q denoting the $q\times q$ identity matrix. Hence, given $\hat{H}$ in (2.10),

$$n\hat{d}_p'\hat{H}^{11}\hat{d}_p = n\tilde{d}_U(0_{q\times 1},\beta_0)'\hat{\mathcal{V}}\tilde{d}_U(0_{q\times 1},\ \beta_0), \tag{C.34}$$

with

$$\hat{\mathcal{V}} = \begin{pmatrix} I_q \\ -\hat{D}_{22}^{-1}\hat{D}_{21}\end{pmatrix}\hat{H}^{11}\left(I_q\ ;\ -\hat{D}_{12}\hat{D}_{22}^{-1}\right). \tag{C.35}$$

Thus,

$$n\hat{d}'\hat{H}^{-1}\hat{d} = n\hat{d}_p'\hat{H}^{11}\hat{d}_p = n\tilde{d}_U(0_{q\times 1},\ \beta_0)'\hat{\mathcal{V}}\tilde{d}_U(0_{q\times 1},\ \beta_0) \tag{C.36}$$

However, under $\mathcal{H}_{1A}$, $\tilde{d}_U(0_{q\times 1},\ \beta_0)$ is no longer evaluated at the true parameter value as $\mu_0 \neq 0$. By MVT around μ_0, we can write

$$\tilde{d}_U(0_{q\times 1},\ \beta_0) = \tilde{d}_U(\mu_0,\ \beta_0) - \frac{\partial \tilde{d}_U(\bar{\mu},\ \beta_0)}{\partial \mu}\mu_0 \equiv \tilde{d}_U(\mu_0,\ \beta_0) - \tau, \tag{C.37}$$

with $\bar{\mu}$ being intermediate point such that $\|\bar{\mu}-\mu_0\| \leq \|\mu_0\|$ and τ being the $q+k\times 1$ vector defined as

$$\tau = \frac{\partial \tilde{d}_U(\bar{\mu},\ \beta_0)}{\partial \mu}\mu_0 = \frac{2}{n}U'\mathcal{P}_Z\Upsilon_{f,g}\mu_0 = \hat{J}'\hat{M}^{-1}\hat{\Xi}\mu_0. \tag{C.38}$$

Similarly to (C.2) and (C.11),

$$\|\tilde{d}_U(\mu_0,\ \beta_0)\| \leq K\|\hat{J}\|\|\hat{M}^{-1}\| \left\|\frac{1}{n}Z'\epsilon\right\| + K\|\hat{J}\|\|\hat{M}^{-1}\| \left\|\frac{1}{n}Z'R\right\|$$
$$= O_p\left(\max\left(\sqrt{\frac{p}{n}}, p^{-\nu}\right)\right) = O_p\left(\sqrt{\frac{p}{n}}\right) \tag{C.39}$$

for ν satisfying $\sqrt{n}/p^{\nu+1/2} = o(1)$, which holds under Assumption 3, and $\|\tau\| = O_p(1)$ and non-zero, since $\mu_0 \neq 0$.

We furthermore define the unconstrained version of $\hat{H}$ evaluated at generic parameters' value as

$$\tilde{H}_U(\mu,\beta) = 4\hat{J}'\hat{M}^{-1}\tilde{\Omega}_U(\mu,\beta)\hat{M}^{-1}\hat{J}, \tag{C.40}$$

partitioned in the usual way, where $\tilde{\Omega}_U$ is defined according to (3.12). We also define its limit quantity $H_U(\mu_0,\ \beta_0) = 4J'M^{-1}\Omega M^{-1}J$, where, as previously defined, $\Omega = n^{-1}\mathbb{E}(Z'\Sigma Z)$ and Σ is the $n \times n$ block-diagonal matrix with $n_g \times n_g$ diagonal block given by $\Sigma_g, g = 1,\ldots,G$. Similar to earlier calculations in the proof of Theorem C2, under Assumptions 3-5, $\|\tilde{H}_U(\mu,\beta)\| = O_p(1)$, uniformly in (μ,β) and $\liminf_{n\to\infty} \underline{eig}(\tilde{H}_U(\mu,\beta)) > c > 0$, uniformly in (μ,β) and almost surely.

Clearly, $\hat{H} = \tilde{H}_U(0,\hat{\beta})$. We can apply the MVT to $\hat{H}^{-1}$ around the true parameters' value and obtain

$$\hat{H}^{-1} = \tilde{H}_U^{-1}(\mu_0,\beta_0) + \sum_{j=1}^{p} \tilde{H}_U^{-1}(\bar{\mu},\bar{\beta})\frac{\partial \tilde{H}_U}{\partial \mu_j}|_{(\bar{\mu},\bar{\beta})}\tilde{H}_U^{-1}(\bar{\mu},\bar{\beta})\mu_{0j}$$
$$- \sum_{t=1}^{k} \tilde{H}_U^{-1}(\bar{\mu},\bar{\beta})\frac{\partial \tilde{H}_U}{\partial \beta_t}|_{(\bar{\mu},\bar{\beta})}\tilde{H}_U^{-1}(\bar{\mu},\bar{\beta})(\hat{\beta}_t - \beta_{0t}) \equiv \tilde{H}_U^{-1}(\mu_0,\beta_0) + T, \tag{C.41}$$

where $\bar{\mu}$ and $\bar{\beta}$ are intermediate points such that $\|\bar{\mu}-\mu_0\| \leq \|\mu_0\|$ and $\|\bar{\beta}-\beta_0\| \leq \|\hat{\beta}-\beta_0\|$. Under $\mathcal{H}_{0A}$, $\|T\| = O_p(\sqrt{p/n})$. Under $\mathcal{H}_{1A}$, $\mu_{0j} \neq 0$ for some $j = 1,\ldots,p$ and, since $\hat{\beta}_t$ for $t = 1,\ldots,k$ are restricted estimates, $\hat{\beta}_t - \beta_{0t} = O_p(1)$ for some $t = 1,\ldots,k$. Thus, under Assumptions 3-5, $\|T\| = O_p(p)$ and $\liminf_{n\to\infty} \underline{eig}(T) > c > 0$. By partitioning T in the usual way, we obtain $\hat{H}^{11} = \tilde{H}_U^{11}(\mu_0,\ \beta_0) + T_{11}$. Also, let

$$\tilde{\mathcal{V}}(\mu_0,\ \beta_0) = \begin{pmatrix} I_q \\ -\hat{D}_{22}^{-1}\hat{D}_{21} \end{pmatrix} \tilde{H}_U^{11}(\mu_0,\ \beta_0)\left(I_q\ ;\ -\hat{D}_{12}\hat{D}_{22}^{-1}\right) \tag{C.42}$$

and

$$\tilde{\mathcal{W}} = \begin{pmatrix} I_q \\ -\hat{D}_{22}^{-1}\hat{D}_{21} \end{pmatrix} T_{11} \left(I_q \ ; \ -\hat{D}_{12}\hat{D}_{22}^{-1} \right). \tag{C.43}$$

From (C.37) and (C.41), (C.36) becomes

$$\begin{aligned} n\hat{d}_p'\hat{H}^{11}\hat{d}_p =& n\tilde{d}_U(\mu_0,\ \beta_0)'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tilde{d}_U(\mu_0,\ \beta_0) + 2n\tau'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tilde{d}_U(\mu_0,\ \beta_0) \\ &+ n\tau'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tau + n\tilde{d}_U(\mu_0,\ \beta_0)'\tilde{\mathcal{W}}\tilde{d}_U(\mu_0,\ \beta_0) \\ &+ 2n\tau'\tilde{\mathcal{W}}\tilde{d}_U(\mu_0,\ \beta_0) + n\tau'\tilde{\mathcal{W}}\tau, \end{aligned} \tag{C.44}$$

and thus

$$\begin{aligned} \frac{n\hat{d}_p'\hat{H}^{11}\hat{d}_p - q}{(2q)^{1/2}} =& \frac{n\tilde{d}_U(\mu_0,\ \beta_0)'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tilde{d}_U(\mu_0,\ \beta_0) - q}{(2q)^{1/2}} \\ &+ \frac{\sqrt{2}n}{\sqrt{q}}\tau'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tilde{d}_U(\mu_0,\ \beta_0) \\ &+ \frac{n}{\sqrt{2q}}\tau'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tau + \frac{n}{\sqrt{2q}}\tilde{d}_U(\mu_0,\ \beta_0)'\tilde{\mathcal{W}}\tilde{d}_U(\mu_0,\ \beta_0) \\ &+ \frac{\sqrt{2}n}{\sqrt{q}}\tau'\tilde{\mathcal{W}}\tilde{d}_U(\mu_0,\ \beta_0) + \frac{n}{\sqrt{2q}}\tau'\tilde{\mathcal{W}}\tau \end{aligned} \tag{C.45}$$

By a similar argument adopted in the proof of Theorem C1, we can show $\|\tilde{d}_U(\mu_0,\ \beta_0) - d_U\| = O_p(p^{3/2}/n)$, with $d_U = -2/nJ'M^{-1}Z'\epsilon$ and $d_p = (I_q; -D_{12}D_{22}^{-1})d_U$. Also, we can show

$$\|\tilde{H}_U(\mu_0,\ \beta_0) - H_U\| = O_p\left(\frac{p}{\sqrt{n}}\right), \tag{C.46}$$

such that, under Assumptions 3-5, $\|\tilde{H}_U^{11}(\mu_0,\ \beta_0) - H_U^{11}\| = O_p\left(p/\sqrt{n}\right)$. We show the claim in (C.46) by routine arguments as in (C.19) and (C.20), after observing that $\tilde{H}_U(\mu_0,\ \beta_0) = 4\hat{J}'\hat{M}^{-1}\tilde{\Phi}_R\hat{M}^{-1}\hat{J}$, with $\tilde{\Phi}_R = \tilde{\Phi} + \sum_{i=1}^n z_i z_i' R_i^2/n$, and

$$\begin{aligned} &\left\|\Delta_{\Phi}^{\tilde{\Phi}_R}\right\| \le \left\|\Delta_{\Phi}^{\tilde{\Phi}}\right\| + \left\|\frac{\sum_{i=1}^n z_i z_i' R_i^2}{n}\right\| = O_p\left(\frac{p}{\sqrt{n}}\right) + \sup_{1\le i\le n} R_i^2\ \|\hat{M}\| \\ =& O_p\left(\frac{p}{\sqrt{n}}\right) + O_p(p^{-2\nu}) = O_p\left(\frac{p}{\sqrt{n}}\right), \end{aligned} \tag{C.47}$$

where the last equality follows for ν satisfying $\sqrt{n}/p^{2\nu+1} = o(1)$, which holds under Assumption 3.

After showing, similarly to what done in the proof of Theorem C2, that

$$\tilde{d}_U(\mu_0,\ \beta_0)'\tilde{\mathcal{V}}(\mu_0,\ \beta_0)\tilde{d}_U(\mu_0,\ \beta_0) - d_p' H_U^{11} d_p = o_p\left(\frac{\sqrt{p}}{n}\right), \tag{C.48}$$

we conclude that the first term in (C.45) is $O_p(1)$, as shown in Theorem 1. By standard norm inequalities, the second term in (C.45) is $O_p(\sqrt{n})$, the third is $O_p(n/\sqrt{p})$, the fourth is $O_p(p^{3/2})$, the fifth is $O_p(p\sqrt{n})$ and the sixth is $O_p(n\sqrt{p})$. The last term dominates the former five ones and thus, under $\mathcal{H}_{1A}$, for all $\eta > 0$, $\mathbb{P}\left(|\mathcal{S}|^{-1} \leq \eta/n\sqrt{p}\right) \to 1$ as $n \to \infty$ and hence consistency of $\mathcal{S}$ follows. □

D Auxiliary lemmas

Lemma D1. *Let $p^2/n \to 0$ as $n \to \infty$ and suppose that Assumptions 2-4 hold with $\nu > 3/2$. Then, as $n \to \infty$,*

$$\left\|\hat{M} - M\right\| = O_p\left(\frac{p}{\sqrt{n}}\right), \left\|\hat{J} - J\right\| = O_p\left(\frac{p}{\sqrt{n}}\right). \tag{D.1}$$

Proof of Lemma D1: This is Lemma 1 in Gupta et al. (2025). □

Lemma D2. *Under Assumptions 1 and 3,*

$$\limsup_{n\to\infty} \overline{eig}(\Phi) < \infty \text{ and } \liminf_{n\to\infty} \underline{eig}(\Phi) > 0.$$

Proof. Let x be a non-stochastic $m \times 1$ vector with $\|x\| = 1$. Then

$$x'\Phi x = \mathbb{E}\left(x'n^{-1}Z'\Sigma Z x\right) \leq \mathbb{E}\left(x'n^{-1}Z'Z\overline{eig}(\Sigma)x\right) = (x'Mx)\overline{eig}(\Sigma) \leq \overline{eig}(M)\overline{eig}(\Sigma),$$

uniformly over x such that $\|x\| = 1$. Then the claim for $\overline{eig}(\Phi)$ follows by (3.1) and (3.2). The proof of the claim for $\underline{eig}(\Phi)$ is similar. □

E Original results

In this section, we reproduce and detail the original empirical findings from Section 4. These results serve as a reference for comparison with our nonparametric test results, allowing readers to assess how our procedure extends the original analyses. Here we retain

Roman numeral column numbering to differentiate these tables from the test tables in the main text and to facilitate cross-referencing.

E.1 Professional golf tournaments (Guryan et al., 2009)

Table E1 reproduces the estimates in Guryan et al. (2009), who study peer effects in professional golf tournaments. Column (i) reports the authors' baseline regression results from estimating the following equation:

$$y_{i,tr} = \alpha + \beta Ability_i + \gamma\, w'_{i,tr} Ability + \delta_{tc} + \varepsilon_{i,tr}, \tag{E.1}$$

where $y_{i,tr}$ denotes the score of player i in group k, round r, and tournament t. Player performance depends on own ability, $Ability_i$, measured by the corrected handicap score, and on peer ability, summarized by the covariate social exposure term $w'_{i,tr}Ability$. Let w denote a peer-weight matrix defined over the stacked player–group–round–tournament observations, with generic element $w_{i,tr,j}$ indicating the relevance of player j for player i within the same playing group at round r, and tournament t. The vector $w_{i,tr}$ denotes the row of w corresponding to observation (i, r, t) and collects the weights assigned to player i for all other players in the same group. *Ability* is a vector compiling the average handicap measure of each player over the last 2-3 years.[16] Accordingly, $w'_{i,tr}Ability = \sum_{j \neq i} w_{i,tr,j} Ability_{t,j}$ denotes the weighted mean ability of peers faced by player i in a given round and tournament. The coefficients β and γ capture the effects of individual and peer ability on playing scores, respectively. The term δ_{tc} denotes tournament-by-category fixed effects, and $\varepsilon_{i,tr}$ is an idiosyncratic error term.

Column (ii) considers alternative measures of playing partners' ability that may influence performance through different channels. Specifically, they replace partners' handicap with measures such as *driving distance*, number of *putts*, and *greens* hit in regulation, which help distinguish potential "learning" effects from pure "motivation" effects.[17] In Column (iii) the authors also include interaction terms allowing peer effects to vary with a player's own ability (measured on the basis of the corrected handicap) and experience.

[16] To construct the individual *Ability* variable, the authors use scores from the previous three years for 2002 data and scores from the previous two years for 2005 and 2006 data.

[17] The intuition is that players may learn about wind or course conditions from observing another player's putting, but cannot directly learn to drive longer; the driving-distance coefficient thus captures the motivation component net of learning.

Table E1: Original results by Guryan et al. (2009)

Dependent var.	(i)	(ii)	(iii)
$Ability_i$	0.672^{***}		0.656^{***}
	(0.039)		(0.039)
$w'_{i,tr}Ability$	-0.035		-0.036
	(0.040)		(0.040)
$DrivDist_i$		-0.009	
		(0.004)	
$w'_{i,tr}DrivDist$		0.003	
		(0.004)	
$Putts_i$		0.130^{***}	
		(0.030)	
$w'_{i,tr}Putts$		-0.045	
		(0.039)	
$Greens_i$		-0.682^{***}	
		(0.050)	
$w'_{i,tr}Greens_t$		-0.023	
		(0.060)	
$Ability_i \times w'_{i,tr}Ability$			0.081
			(0.033)
Exp_{it}			0.019^{***}
			(0.004)
$Exp_{it} \times w'_{i,tr}Ability$			0.015^{**}
			(0.005)
Tournament × category fixed effects	✓	✓	✓
n	17,492	17,182	17,492

Notes: Regression results in Columns (i) and (ii) replicate Columns (1) and (5) from Table 5 in Guryan et al. (2009). Column (iii) replicates Column (4) from Table 8 in Guryan et al. (2009). The dependent variable is the golf score of player i in a given round. $Ability_i$ denotes player i's ability, measured by the average handicap score of the last 2-3 years. $w'_{i,tr}Ability = \sum_{j \neq i} w_{i,tr,j} Ability_j$ denotes the weighted average ability of player i's peers j. Column (ii) includes additional ability measures such as driving distance ($DrivDist_i$), number of putts ($Putts_i$), and greens in regulation ($Greens_i$), along with their peer analogues $w'_{i,tr}DrivDist$, $w'_{i,tr}Putts$, and $w'_{i,tr}Greens$, constructed analogously using the weights in $w_{i,tr}$. Column (iii) further includes interactions between own ability and peer ability, $Ability_i \times w'_{i,tr}Ability$, and between experience and peer ability, $Exp_i \times w'_{i,tr}Ability$, where Exp_i denotes player experience. All specifications include tournament-by-category fixed effects. Standard errors clustered at the playing-group level are reported in parentheses. Observations are weighted by the inverse of the sample variance of the ability measure, following Guryan et al. (2009). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Overall, these results based on linear specifications find limited evidence of peer effects in individual performance. None of the peer coefficients in specifications (i) and (ii) are statistically significant, while in (iii) some evidence of heterogeneity emerges. Interacting average peer ability with own experience yields a small but statistically significant effect: for a player with one year of experience, a one-stroke increase in average partner ability decreases the player's score by about 0.02 strokes[18], indicating sensitivity to competitive pressure early in the career. As experience increases, players become progressively less sensitive: the initially negative peer effect fades out around 2–3 years of experience and eventually reverses, so that more experienced players perform better in stronger competitive fields and worse in weaker ones.

E.2 Network expansion and firm performance (Cai and Szeidl, 2017)

In what follows, we replicate the peer-quality specifications of Cai and Szeidl (2017), which relate firm performance to the baseline employment size of other firms in the same meeting group. Specifically, Table 8 in Cai and Szeidl (2017) estimates the following model on the sample of treated firms:

$$y_{i,t} = \alpha + \delta Post_{i,t} + \gamma \, w'_{i,t} EmpSize + x'_i \beta + \theta_i + \varepsilon_{i,t}, \tag{E.2}$$

where $y_{i,t}$ denotes a generic measure of firm i's performance at time t. The vector $w_{i,t}$ denotes the peer-weight vector associated with observation (i, t) and collects the weights assigned to other firms in the same meeting group. Let w denote the peer-weight matrix defined over the stacked panel of firm–time observations, so that $w_{i,t}$ corresponds to the row of w associated with firm i at time t. The vector $EmpSize$ collects firms' baseline log employment sizes, measured prior to the intervention, such that $w'_{i,t} EmpSize = \sum_{j \neq i} w_{i,t,j} EmpSize_j$ captures the average baseline employment size of firm i's peers at time t. The term x_i collects firm-level control variables (including size category, sector, subregion, and their interactions); θ_i are firm fixed effects; and $\varepsilon_{i,t}$ is an idiosyncratic error term.

Panel A of Table E2 focuses on standard performance outcomes. Column (i) shows that larger peers are associated with higher sales: the estimated coefficient implies that being randomized into a group with peers with a 10% increase in average peer size raises log sales by about 1%. Column (ii) also finds a statistically significant effect on profits (around

[18] A back of the envelope calculation using the coefficients in column (iii) of Table E1.

RMB 27,800). In contrast, the estimates for employment, total assets, and productivity are non-significant. Materials and utility costs (v and vi), however, respond positively and with magnitudes comparable to the sales effect, consistent with higher scale of operations rather than changes in measured productivity.

Panel B turns to intermediate outcomes and potential alternative channels. Here, log peer size is positively and significantly related to the number of clients and to the management score (viii and xi), while effects on the number of suppliers and bank loans are small and not statistically significant. The final two columns—differences between reported and book sales, and the tax-to-sales ratio—show coefficients close to zero, which is reassuring for concerns about differential misreporting or tax evasion. Overall, the pattern in Table E2 suggests that being matched with larger peers raises firm scale and improves some management practices, without clear evidence of changes in tax behavior or accounting practices.

Table E2: Original results by Cai and Szeidl (2017)

	Panel A: Main performance measures						
Dependent var.	log Sales (i)	Profit (10,000 RMB) (ii)	log Number of employees (iii)	log Total assets (iv)	log Material cost (v)	log Utility cost (vi)	log Productivity (vii)
$w'_{i,t}EmpSize$	0.105***	27.825**	0.043	-0.016	0.100*	0.141***	0.029
	(0.040)	(13.432)	(0.032)	(0.034)	(0.052)	(0.042)	(0.020)
Firm demographics	✓	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓	✓	✓	✓	✓
n	4,183	4,076	4,183	4,183	4,148	4,086	4,183

	Panel B: Intermediate outcomes and alternative explanations						
Dependent var.	log Number of clients (viii)	log Number of suppliers (ix)	Bank loan (x)	Management (xi)	Innovation (xii)	log Reported - log book sales (xiii)	Tax/sales (xiv)
$w'_{i,t}EmpSize$	0.068**	-0.001	0.017	0.162***	0.027	0.022	-0.001
	(0.032)	(0.030)	(0.016)	(0.027)	(0.017)	(0.014)	(0.001)
Firm demographics	✓	✓	✓	✓	✓	✓	✓
Firm fixed effects	✓	✓	✓			✓	✓
n	4,173	4,170	4,183	2,774	1,409	4,152	4,178

Note: Results originally presented in Table 8 from Cai and Szeidl (2017). Regressions only use data for treated firms. The term $w'_{i,t}EmpSize$ is the average baseline log employment of other group members. Firm demographics are size category, sector, subregion, and their interactions. Standard errors clustered at the meeting group level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

E.3 Student achievement (Booij et al., 2017)

Table 4 of the paper by Booij et al. (2017) reports five regression specifications of growing complexity, linking peer-group composition to first-year credit completion. These results, which serve us as basis for the tests in Table 3, are reproduced here below. For exposition, we discuss the fully saturated specification of column (v) which nests all other models. Precisely, Equation (E.3) allows peer effects to operate both directly, through $\mathbf{z}_i'\boldsymbol{\gamma}$, and heterogeneously with respect to own ability, through the interaction $GPA_i\,\mathbf{z}_i'\boldsymbol{\eta}$.

$$\begin{aligned}\mathbf{z}_i &\equiv \big(w_{i,\text{avg}}'GPA,\ w_{i,\text{sd}}'GPA,\ (w_{i,\text{avg}}'GPA)\times(w_{i,\text{sd}}'GPA)\big)'.\\ y_{ig} &= \alpha + \mathbf{z}_i'\boldsymbol{\gamma} + GPA_i\,\mathbf{z}_i'\boldsymbol{\eta} + x_{ig}'\beta + \varepsilon_{ig}.\end{aligned} \tag{E.3}$$

where y_{ig} denotes the academic outcome of student i in group g, measured by the number of credit points obtained during the academic year. Peer academic exposure is summarized by the vector $\mathbf{z}_i \equiv \big(w_{i,\text{avg}}'GPA,\ w_{i,\text{sd}}'GPA,\ (w_{i,\text{avg}}'GPA)\times(w_{i,\text{sd}}'GPA)\big)'$ where w denotes a peer-weight matrix with generic element w_{ij} indicating the relevance of student j for student i within group g. The vector w_i denotes the ith row of w and collects the weights assigned by student i to all other students in the group. Accordingly, $w_{i,\text{avg}}'GPA$ and $w_{i,\text{sd}}'GPA$ denote, respectively, the mean and standard deviation of peers' GPA, computed using the weights in w_i. The coefficient vector $\boldsymbol{\gamma}$ captures the direct association between peer academic characteristics and student outcomes, while $\boldsymbol{\eta}$ allows these associations to vary with the student's own prior academic performance, GPA_i. The vector x_{ig} contains additional individual- and group-level control variables, α is a constant term, and ε_{ig} denotes an idiosyncratic error term. The term x_i includes randomization controls (GPA category, math track, cohort, application order) alone or combined with student demographics (gender, age, professional college attendance) and individual GPA.

Table E3 presents the results from five progressively augmented specifications linking peer-group composition to first-year credit completion. Columns (i)–(iii) include only main effects of peer GPA moments and yield statistically insignificant estimates. Once the interaction between peer mean GPA and dispersion is introduced in column (iv), both mean peer GPA and the mean–SD interaction become statistically significant, while the dispersion term remains negative. Finally, column (v) includes high-order effects allowing peer effects to vary with students' own GPA, and again concludes for statistically significant effects.

Table E3: Original results by Booij et al. (2017)

Dependent var.	(i)	(ii)	(iii)	(iv)	(v)
$w'_{i,\text{avg}}GPA$	0.051	0.048	0.070	0.095**	0.148***
	(0.043)	(0.041)	(0.043)	(0.046)	(0.052)
$w'_{i,\text{sd}}GPA$			-0.095	-0.121*	-0.185**
			(0.073)	(0.063)	(0.082)
$(w'_{i,\text{avg}}GPA) \times (w'_{i,\text{sd}}GPA)$				0.423**	0.343*
				(0.176)	(0.190)
$GPA_i \times w'_{i,\text{avg}}GPA$					-0.117***
					(0.042)
$GPA_i \times w'_{i,\text{sd}}GPA$					0.104
					(0.075)
$GPA_i \times (w'_{i,\text{avg}}GPA) \times (w'_{i,\text{sd}}GPA)$					-0.287**
					(0.138)
Controls					
Randomization	✓	✓	✓	✓	✓
Background		✓	✓	✓	✓
GPA_i		✓	✓	✓	✓

Note: Results originally presented in Table 4 of Booij et al. (2017). The dependent variable is the number of credit points collected by student i in the first academic year. The term $w_i^{\text{avg}}GPA$ denotes the mean GPA of student i's peers, and $w'_{i,\text{sd}}GPA$ denotes standard deviation of peers' GPA. The interaction terms capture complementarities between peer mean GPA and peer GPA dispersion, as well as heterogeneity with respect to student i's own prior academic performance, GPA_i. All regressions include randomization controls (a saturated set of own GPA category, advanced math track, and cohort dummies interacted with application order). Background controls include gender, age, and an indicator for professional college attendance. Standard errors clustered at the tutorial-group level are reported in parentheses. $n = 1{,}876$. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.