SUPPLEMENT TO "FIXED-EFFECT REGRESSIONS ON NETWORK DATA"

(*Econometrica*, Vol. 87, No. 5, September 2019, 1543–1560)

KOEN JOCHMANS
Faculty of Economics, University of Cambridge

MARTIN WEIDNER
Department of Economics, University College London

S.1. ADDITIONAL ILLUSTRATIONS

RECALL THAT OUR MEASURE OF GLOBAL CONNECTIVITY OF THE GRAPH $\mathcal{G}$ is λ_2, the second smallest eigenvalue of the normalized Laplacian matrix. In the following discussion, we provide some concrete examples of graphs for which λ_2 can be explicitly calculated, and we discuss the implications of our variance bound in Theorem 2.

Our first example illustrates that even if $\lambda_2 \to 0$ with the sample size, we may still have that $\text{var}(\hat{\alpha}_i) \asymp d_i^{-1}$.

EXAMPLE S.1—Hypercube graph: Consider the N-dimensional hypercube, where each of $n = 2^N$ vertices is involved in N edges; see the left-hand side of Figure S.1. This is an N-regular graph—that is, $d_i = h_i = N$ for all i—with the total number of edges in the graph equalling 2^{N-1}. Here,

$$\lambda_2 = \frac{2}{N} = O\big((\ln n)^{-1}\big).$$

Thus, $\lambda_2 h_i$ is constant in n. An application of Theorem 2 yields

$$1 + o(1) \leq \frac{N \, \text{var}(\hat{\alpha}_i)}{\sigma^2} \leq \frac{3}{2} + o(1).$$

From this, we obtain the convergence rate result $(\hat{\alpha}_i - \alpha_i) = O_p((\ln n)^{-1/2})$.

Theorem 2 allows us to establish the convergence rate for the hypercube, but the conditions are too stringent to obtain (12). The reason is that h_i does not increase fast enough to ensure that $\lambda_2 h_i \to \infty$. The following example deals with an extended hypercube and illustrates that, despite $\lambda_2 \to 0$, we still have $\lambda_2 h_i \to \infty$ in this case.

EXAMPLE S.2—Extended hypercube graph: Start with the N-dimensional hypercube $\mathcal{G}$ from the previous example and add edges between all path-2 neighbors in $\mathcal{G}$; see the right-hand side of Figure S.1 for an example. The resulting graph still has $n = 2^N$ vertices, but now has $N(N+1)2^{N-1}$ edges. Here

$$d_i = h_i = \frac{N(N+1)}{2}, \qquad \lambda_2 = \frac{4}{N+1},$$

so that $\lambda_2 h_i \to \infty$ holds, despite $\lambda_2 \to 0$ as $n \to \infty$. Theorem 2 therefore implies (12) in this example.

Koen Jochmans: kj345@cam.ac.uk
Martin Weidner: m.weidner@ucl.ac.uk

 https://doi.org/10.3982/ECTA14605

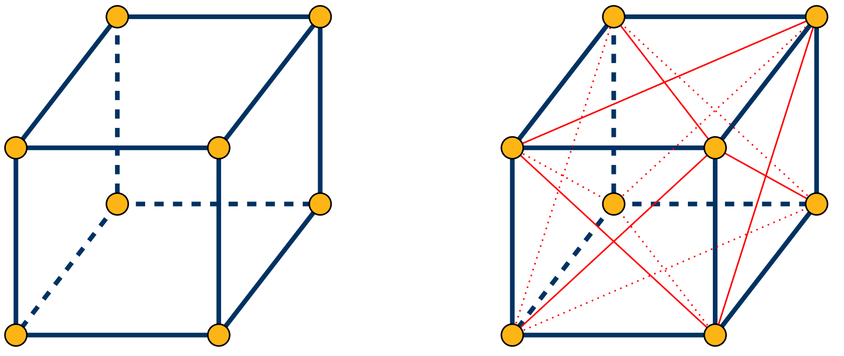

FIGURE S.1.—Three-dimensional hypercube (left) and extended hypercube (right).

The next example shows that our bound can still be informative if h_i is finite.

EXAMPLE S.3—Star graph: Consider a star graph around the central vertex 1, that is, the graph with n vertices and edges

$$E = \{(1, j) : 2 \leq j \leq n\};$$

see the left-hand side of Figure S.2. Here, $\lambda_2 = 1$ for any n, while $d_1 = n - 1$, $h_1 = 1$ and $d_i = 1$, $h_i = n - 1$ for $i \neq 1$. For $i = 1$, one finds that the bounds in Theorem 2 imply that $\mathrm{var}(\hat{\alpha}_1) = O(n^{-1})$, and so

$$(\hat{\alpha}_1 - \alpha_1) = O_p(n^{-1/2}).$$

In contrast, for $i \neq 1$, we find $\lambda_2 h_i \to \infty$ and thus, although (12) holds, these α_i cannot be estimated consistently as $d_i = 1$.

The previous example also illustrates that λ_2 can be large despite having many vertices with small degrees. It is largely due to this property that we prefer to measure global connectivity by λ_2 and not by the "algebraic connectivity" (the second smallest eigenvalue of $\boldsymbol{L}$; see, e.g., Chung 1997), which has been studied more extensively.

Our last example shows the effect on the upper bound in Theorem 2 when neighbors themselves are more strongly connected.

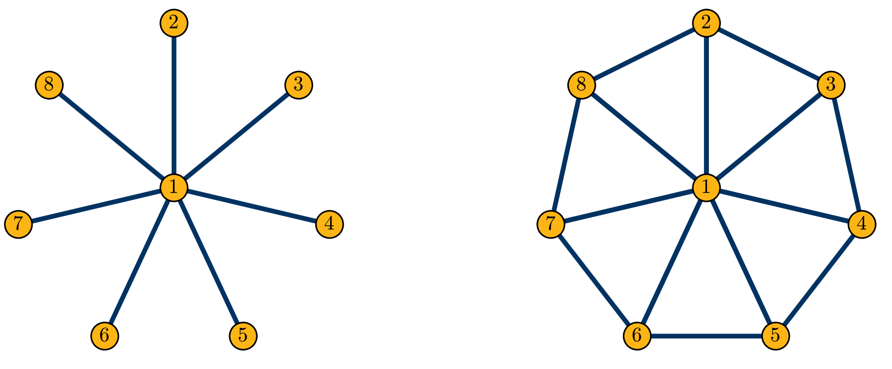

FIGURE S.2.—Star graph (left) and wheel graph (right) for $n = 8$.

EXAMPLE S.4—Wheel graph: The wheel graph is obtained by combining a star graph centered at vertex 1 with a cycle graph on the remaining $n-1$ vertices; see the right-hand side of Figure S.2. Thus, a wheel graph contains strictly more edges than the underlying star graph, although none of these involves the central vertex directly. From Butler (2016), we have

$$\lambda_2 = \min\left\{\frac{4}{3}, 1 - \frac{2}{3}\cos\left(\frac{2\pi}{n}\right)\right\},$$

which satisfies $\lambda_2 \geq 1$ only for $n \leq 4$, and converges to $1/3$ at an exponential rate. However, while, as in the star graph, $d_1 = n-1$, we now have that $h_i = 3$ for all $i \neq 1$. Hence, $\lambda_2 h_1 > 1$ for any finite n and the upper bound in Theorem 2 is strictly smaller than in the star graph.

The last two examples also illustrate that adding edges to the graph (in this case, to obtain the wheel graph from the star graph) can result in a decrease of our measure of global connectivity λ_2. This is not a problem, however, for our results, as we only require that λ_2 be sufficiently different from zero. The wheel graph with $\lambda_2 \geq 1/3$, for example, clearly describes a very well globally connected graph by that measure.

S.2. VARIANCE BOUNDS FOR DIFFERENCES

Our focus in the main text has been inference on the α_i, under the constraint in (3), $\sum_i d_i \alpha_i = 0$. An alternative to normalizing the parameters that may be useful in certain applications is to focus directly on the differences $\alpha_i - \alpha_j$ for all $i \neq j$. An example where this is the case is Finkelstein, Gentzkow, and Williams (2016). We give a corresponding version of Theorem 2 here.

Let $d_{ij} := \sum_{k \in V} (\boldsymbol{A})_{ik} (\boldsymbol{A})_{jk}$ for an unweighted graph $d_{ij} = |[i] \cap [j]|$, the number of vertices that are neighbors of both i and j. Write

$$h_{ij} := \begin{cases} \left(\dfrac{1}{d_{ij}} \displaystyle\sum_{k \in V} \dfrac{(\boldsymbol{A})_{ik} (\boldsymbol{A})_{jk}}{d_k}\right)^{-1} & \text{for } d_{ij} \neq 0, \\ \infty & \text{for } d_{ij} = 0, \end{cases}$$

for the corresponding harmonic mean of the degrees of the vertices $k \in [i] \cap [j]$. We have the following theorem.

THEOREM S.1—First-order bound for differences: *Let $\mathcal{G}$ be connected. Then*

$$\sigma^2\left(\frac{1}{d_i} + \frac{1}{d_j} - \frac{2(\boldsymbol{A})_{ij}}{d_i d_j}\right) \leq \operatorname{var}(\hat{\alpha}_i - \hat{\alpha}_j) \leq \sigma^2\left(\frac{1}{d_i} + \frac{1}{d_j} - \frac{2(\boldsymbol{A})_{ij}}{d_i d_j}\right) + \frac{\sigma^2}{\lambda_2}\left(\frac{1}{d_i h_i} + \frac{1}{d_j h_j} - \frac{2 d_{ij}}{d_i d_j h_{ij}}\right).$$

For a simple graph $\mathcal{G}$, when $[i] = [j]$ but $i \notin [j]$ and $i \notin [j]$, that is, when vertices i and j share exactly the same neighbors and are not connected themselves, the theorem implies

$$\operatorname{var}(\hat{\alpha}_i - \hat{\alpha}_j) = \sigma^2\left(\frac{1}{d_i} + \frac{1}{d_j}\right), \tag{S.1}$$

as, in that case, both $(\boldsymbol{A})_{ij}$ and the second term in the upper bound in Theorem S.1 are zero.

S.3. ALTERNATIVE NORMALIZATION

If we change the normalization constraint in the least-squares minimization problem (4) to

$$\sum_{i=1}^{n} \alpha_i = 0,$$

we obtain the estimator $\hat{\boldsymbol{\alpha}}^{\diamond} = \boldsymbol{M}_{\boldsymbol{\iota}} \hat{\boldsymbol{\alpha}}$, where $\boldsymbol{M}_{\boldsymbol{\iota}} = \boldsymbol{I}_n - n^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'$ is the projector orthogonal to $\boldsymbol{\iota}_n$. We then have $\operatorname{var}(\hat{\boldsymbol{\alpha}}^{\diamond}) = \sigma^2 \boldsymbol{L}^{\dagger}$, because this variance needs to satisfy $\operatorname{var}(\hat{\boldsymbol{\alpha}}^{\diamond})\boldsymbol{\iota}_n = 0$, and the Moore–Penrose pseudoinverse guarantees that the null space of $\boldsymbol{L}$ equals the null space of $\boldsymbol{L}^{\dagger}$. Thus, changing the normalization corresponds to changing the particular pseudoinverse of $\boldsymbol{L}$ that features in the expression for the variance. From $\hat{\boldsymbol{\alpha}}^{\diamond} = \boldsymbol{M}_{\boldsymbol{\iota}} \hat{\boldsymbol{\alpha}}$, we find

$$\operatorname{var}\big(\hat{\boldsymbol{\alpha}}^{\diamond}\big) = \boldsymbol{M}_{\boldsymbol{\iota}} \operatorname{var}(\hat{\boldsymbol{\alpha}}) \boldsymbol{M}_{\boldsymbol{\iota}},$$

which thus also shows that $\boldsymbol{L}^{\dagger} = \boldsymbol{M}_{\boldsymbol{\iota}} \boldsymbol{L}^{\star} \boldsymbol{M}_{\boldsymbol{\iota}}$. We have $\boldsymbol{L}^{\star} \leq \lambda_2^{-1}\boldsymbol{D}^{-1}$, and, therefore, $\boldsymbol{L}^{\dagger} \leq \lambda_2^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}$. We thus find $\operatorname{var}(\hat{\alpha}_i^{\diamond}) = \sigma^2 \boldsymbol{e}_i'\boldsymbol{L}^{\dagger}\boldsymbol{e}_i \leq \lambda_2^{-1}\sigma^2\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i$, and evaluating the last expression gives the following theorem.

THEOREM S.2—Global bound under alternative normalization: *Let $\mathcal{G}$ be connected. Then*

$$\operatorname{var}\big(\hat{\alpha}_i^{\diamond}\big) \leq \frac{1}{d_i}\frac{\sigma^2}{\lambda_2}\left(1 + \frac{d_i}{nh}\right).$$

Notice that $d_i/(nh) \leq 1/h \leq 1$ and, therefore, $\operatorname{var}(\hat{\alpha}_i^{\diamond}) \leq \frac{2}{d_i}\frac{\sigma^2}{\lambda_2}$. For the estimator $\hat{\alpha}_i$ obtained under the normalization in the main text, we immediately find from (6) and $(\boldsymbol{S}^{\dagger})_{ii} \leq \lambda_2^{-1}$ that $\operatorname{var}(\hat{\alpha}_i) \leq \frac{1}{d_i}\frac{\sigma^2}{\lambda_2}$. Thus, for sequences of growing networks, we find the pointwise consistency results $(\hat{\alpha}_i^{\diamond} - \alpha_i) \xrightarrow{p} 0$ and $(\hat{\alpha}_i - \alpha_i) \xrightarrow{p} 0$ for both estimators, under the sufficient condition $\lambda_2 d_i \to \infty$.

Analogously one can extend Theorem 2 from $\hat{\alpha}_i$ to $\hat{\alpha}_i^{\diamond}$ as follows.

THEOREM S.3—First-order bound under alternative normalization: *Let $\mathcal{G}$ be connected. Then*

$$\frac{\sigma^2}{d_i}\left(1 - \frac{2}{n}\right) - \frac{2\sigma^2}{nh_i^{(2)}} \leq \operatorname{var}\big(\hat{\alpha}_i^{\diamond}\big) \leq \frac{\sigma^2}{d_i}\left(1 + \frac{1}{\lambda_2 h_i}\right) + \frac{\sigma^2}{h}\left(\frac{2}{n} + \frac{1}{\lambda_2 H}\right),$$

where $h_i^{(2)} = (\frac{1}{d_i}\sum_{j\in[i]}\frac{(\boldsymbol{A})_{ij}}{d_j})^{-1}$, and h and H are defined in the main text.

Analogous to (12) in the main text, we thus find

$$\operatorname{var}\big(\hat{\alpha}_i^{\diamond}\big) = \frac{\sigma^2}{d_i} + o\big(d_i^{-1}\big),$$

provided that $\lambda_2 h_i \to \infty$ and $nh/d_i \to \infty$ and $nh_i^{(2)}/d_i \to \infty$ and $\lambda_2 hH/d_i \to \infty$ as $n \to \infty$. Therefore, under plausible assumptions on the sequence of growing networks, we find the same asymptotic properties for $\hat{\alpha}_i^{\diamond}$ as for $\hat{\alpha}_i$. The particular choice of normalization in the main text is not necessary for our main results, but it makes all derivations as well as the presentation of the results more convenient.

S.4. PROOFS

PROOF OF THEOREM 1 (EXISTENCE): The estimator is defined by the constraint minimization problem in (4). For convenience, we express the constraint in quadratic form, $(\boldsymbol{a}'\boldsymbol{d})^2 = 0$. By introducing the Lagrange multiplier $\lambda > 0$, we can write

$$\check{\boldsymbol{\alpha}} = \underset{\boldsymbol{a}\in\mathbb{R}^n}{\arg\min}(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{a})'\boldsymbol{M}_X(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{a}) + \lambda\big(\boldsymbol{a}'\boldsymbol{d}\big)^2.$$

Solving the corresponding first-order condition we obtain

$$\begin{aligned}\check{\boldsymbol{\alpha}} &= \big(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B} + \lambda\boldsymbol{d}\boldsymbol{d}'\big)^{-1}\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{y} \\ &= \boldsymbol{D}^{-1/2}\big(\boldsymbol{S}_X + \lambda\boldsymbol{\psi}\boldsymbol{\psi}'\big)^{-1}\boldsymbol{D}^{-1/2}\boldsymbol{B}'\boldsymbol{y},\end{aligned} \tag{S.2}$$

where $\boldsymbol{S}_X := \boldsymbol{D}^{-1/2}\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B}\boldsymbol{D}^{-1/2}$ and $\boldsymbol{\psi} := \boldsymbol{D}^{1/2}\boldsymbol{\iota}_n = \boldsymbol{D}^{-1/2}\boldsymbol{d}$. Since we assume that the graph is connected, we have $d_i > 0$ for all i, that is, $\boldsymbol{D}$ is invertible. Our assumption $\operatorname{rank}((\boldsymbol{X}, \boldsymbol{B})) = p + n - 1$ implies that $\operatorname{rank}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B}) = n - 1$, that is, the zero eigenvalue of $\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B}$ has multiplicity 1. By construction of $\boldsymbol{B}$, we have $\boldsymbol{B}\boldsymbol{\iota}_n = 0$, that is, the zero eigenvector of $\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B}$ is given by $\boldsymbol{\iota}_n$. It follows that the zero eigenvalue $\boldsymbol{S}_X$ has multiplicity 1 and eigenvector $\boldsymbol{\psi}$. This explains why the matrix $\boldsymbol{S}_X + \lambda\boldsymbol{\psi}\boldsymbol{\psi}'$ is invertible, which we already used in (S.2). Furthermore, the matrices $\boldsymbol{S}_X$ and $\boldsymbol{\psi}\boldsymbol{\psi}'$ commute, and by properties of the Moore–Penrose inverse, we thus have

$$\big(\boldsymbol{S}_X + \lambda\boldsymbol{\psi}\boldsymbol{\psi}'\big)^{-1} = \boldsymbol{S}_X^{\dagger} + \lambda^{-1}\big(\boldsymbol{\psi}\boldsymbol{\psi}'\big)^{\dagger}. \tag{S.3}$$

We furthermore have

$$\big(\boldsymbol{\psi}\boldsymbol{\psi}'\big)^{\dagger} = m^{-2}\boldsymbol{\psi}\boldsymbol{\psi}', \tag{S.4}$$

where $m = \boldsymbol{\psi}'\boldsymbol{\psi}$ is the total number of observations. Because $\boldsymbol{B}\boldsymbol{\iota}_n = 0$, the contribution from $(\boldsymbol{\psi}\boldsymbol{\psi}')^{\dagger}$ drops out of (S.2), and we obtain

$$\check{\boldsymbol{\alpha}} = \boldsymbol{D}^{-1/2}\boldsymbol{S}_X^{\dagger}\boldsymbol{D}^{-1/2}\boldsymbol{B}'\boldsymbol{y} = \big(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B}\big)^{\star}\boldsymbol{B}'\boldsymbol{y},$$

according to the definition of the pseudoinverse * in the main text. Notice that $\check{\boldsymbol{\alpha}}$ given in the last display does not depend on λ, and automatically satisfies the constraint $\boldsymbol{d}'\check{\boldsymbol{\alpha}} = 0$, that is, any value of λ can be chosen in the above derivation. *Q.E.D.*

PROOF OF THEOREMS 2 AND S.1 (VARIANCE BOUNDS): We first show that if $\mathcal{G}$ is connected, then

$$0 \leq \big[\operatorname{var}(\hat{\boldsymbol{\alpha}}) - \sigma^2\big(\boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1} - 2m^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'\big)\big] \leq \frac{\sigma^2}{\lambda_2}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}. \tag{S.5}$$

Theorems 2 and S.1 will then follow readily. Analogous to (S.3), we also have $(\boldsymbol{S}+\lambda\boldsymbol{\psi}\boldsymbol{\psi}')^{-1}=\boldsymbol{S}^{\dagger}+\lambda^{-1}(\boldsymbol{\psi}\boldsymbol{\psi}')^{\dagger}$. Using this and (S.4), we find

$$\begin{aligned}\boldsymbol{I}_n &= \big(\boldsymbol{S}+\lambda\boldsymbol{\psi}\boldsymbol{\psi}'\big)^{-1}\big(\boldsymbol{S}+\lambda\boldsymbol{\psi}\boldsymbol{\psi}'\big)\\ &= \big(\boldsymbol{S}^{\dagger}+\lambda^{-1}m^{-2}\boldsymbol{\psi}\boldsymbol{\psi}'\big)\big(\boldsymbol{S}+\lambda\boldsymbol{\psi}\boldsymbol{\psi}'\big),\end{aligned}$$

and since $\boldsymbol{S}\boldsymbol{\psi}=0$ and $\boldsymbol{\psi}'\boldsymbol{\psi}=m$, we thus find that $\boldsymbol{S}^{\dagger}\boldsymbol{S}=\boldsymbol{I}_n-m^{-1}\boldsymbol{\psi}\boldsymbol{\psi}'$, which is simply the idempotent matrix that projects orthogonally to $\boldsymbol{\psi}$. We thus find $\boldsymbol{L}^{\star}\boldsymbol{L}=\boldsymbol{D}^{-1/2}\boldsymbol{S}^{\dagger}\boldsymbol{S}\boldsymbol{D}^{1/2}=\boldsymbol{I}_n-m^{-1}\boldsymbol{\iota}_n\boldsymbol{d}'$. Plugging in $\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{A}$ and then solving for $\boldsymbol{L}^{\star}$ gives

$$\boldsymbol{L}^{\star}=\boldsymbol{D}^{-1}+\boldsymbol{L}^{\star}\boldsymbol{A}\boldsymbol{D}^{-1}-m^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'. \tag{S.6}$$

The Laplacian is symmetric, and so transposition gives

$$\boldsymbol{L}^{\star}=\boldsymbol{D}^{-1}+\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{L}^{\star}-m^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'. \tag{S.7}$$

Replacing $\boldsymbol{L}^{\star}$ on the right-hand side of (S.6) by the expression for $\boldsymbol{L}^{\star}$ given by (S.7), and also using that $\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{\iota}_n=\boldsymbol{\iota}_n$ yields

$$\boldsymbol{L}^{\star}=\boldsymbol{D}^{-1}+\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}+\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{L}^{\star}\boldsymbol{A}\boldsymbol{D}^{-1}-2m^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'. \tag{S.8}$$

Rearranging this equation allows us to write

$$\boldsymbol{L}^{\star}-\big(\boldsymbol{D}^{-1}+\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}-2m^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'\big)=\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{L}^{\star}\boldsymbol{A}\boldsymbol{D}^{-1}.$$

From $\boldsymbol{L}^{*}=\boldsymbol{D}^{-1/2}\boldsymbol{S}^{\dagger}\boldsymbol{D}^{-1/2}$ and $\boldsymbol{0}\le\boldsymbol{S}^{\dagger}\le\lambda_2^{-1}\boldsymbol{I}_n$, we obtain $\boldsymbol{0}\le\boldsymbol{L}^{\star}\le\lambda_2^{-1}\boldsymbol{D}^{-1}$, and therefore,

$$\boldsymbol{0}\le\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{L}^{\star}\boldsymbol{A}\boldsymbol{D}^{-1}\le\lambda_2^{-1}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}.$$

Put together this yields

$$0\le\boldsymbol{L}^{\star}-\big(\boldsymbol{D}^{-1}+\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}-2m^{-1}\boldsymbol{\iota}_n\boldsymbol{\iota}_n'\big)\le\lambda_2^{-1}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1},$$

and multiplication with σ^2 gives the bounds stated in (S.5).

To show Theorems 2 and S.1, we calculate, for $i\neq j$,

$$\begin{aligned}
\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{e}_i&=d_i^{-1}, & \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_i&=d_i^{-1}h_i^{-1},\\
\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{e}_j&=0, & \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_j&=d_i^{-1}d_j^{-1}d_{ij}h_{ij}^{-1},\\
\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_i&=0, & \boldsymbol{e}_i'\boldsymbol{\iota}_n\boldsymbol{\iota}_n'\boldsymbol{e}_i&=1,\\
\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_j&=d_i^{-1}d_j^{-1}(\boldsymbol{A})_{ij}, & \boldsymbol{e}_i'\boldsymbol{\iota}_n\boldsymbol{\iota}_n'\boldsymbol{e}_j&=1,
\end{aligned}$$

where $\boldsymbol{e}_i$ is the vector that has 1 as its ith entry and 0s elsewhere. Combining these results with (S.5) gives the bounds on, respectively, $\operatorname{var}(\hat{\alpha}_i)=\boldsymbol{e}_i'\operatorname{var}(\hat{\boldsymbol{\alpha}})\boldsymbol{e}_i$ and $\operatorname{var}(\hat{\alpha}_i-\hat{\alpha}_j)=(\boldsymbol{e}_i-\boldsymbol{e}_j)'\operatorname{var}(\hat{\boldsymbol{\alpha}})(\boldsymbol{e}_i-\boldsymbol{e}_j)$ stated in the theorems. *Q.E.D.*

PROOF OF THEOREMS S.2 AND S.3: Using that $\boldsymbol{L}^{*}\le\lambda_2^{-1}\boldsymbol{D}^{-1}$ we find that

$$\begin{aligned}\operatorname{var}\big(\hat{\alpha}_i^{\diamond}\big)&=\boldsymbol{e}_i'\operatorname{var}\big(\hat{\boldsymbol{\alpha}}^{\diamond}\big)\boldsymbol{e}_i=\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\operatorname{var}(\hat{\boldsymbol{\alpha}})\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i=\sigma^2\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{L}^{*}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i\\ &\le\lambda_2^{-1}\sigma^2\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i,\end{aligned}$$

and we calculate

$$\begin{aligned} \boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i &= \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{e}_i - \frac{2}{n}\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{\iota}_n + \frac{1}{n^2}\boldsymbol{\iota}_n'\boldsymbol{D}^{-1}\boldsymbol{\iota}_n \\ &= \frac{1}{d_i} - \frac{2}{nd_i} + \frac{1}{nh}. \end{aligned} \tag{S.9}$$

Combing those results gives the statement of Theorem S.2

Next, multiplying $\boldsymbol{M}_{\boldsymbol{\iota}}$ from the left and right to the matrix bounds (S.5) and using $\mathrm{var}(\hat{\boldsymbol{\alpha}}^{\diamond}) = \boldsymbol{M}_{\boldsymbol{\iota}}\,\mathrm{var}(\hat{\boldsymbol{\alpha}})\boldsymbol{M}_{\boldsymbol{\iota}}$ gives

$$0 \leq \left[\mathrm{var}\big(\hat{\boldsymbol{\alpha}}^{\diamond}\big) - \sigma^2\boldsymbol{M}_{\boldsymbol{\iota}}\big(\boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\big)\boldsymbol{M}_{\boldsymbol{\iota}}\right] \leq \frac{\sigma^2}{\lambda_2}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}},$$

and, therefore,

$$0 \leq \left[\mathrm{var}\big(\hat{\alpha}_i^{\diamond}\big) - \sigma^2\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\big(\boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\big)\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i\right] \leq \frac{\sigma^2}{\lambda_2}\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i.$$

We already calculated $\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i$ in (S.9) above. We furthermore have

$$\begin{aligned} \boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i &= \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_i - \frac{2}{n}\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{\iota}_n + \frac{1}{n^2}\boldsymbol{\iota}_n'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{\iota}_n \\ &= 0 - \frac{2}{nd_i}\sum_{j\in[i]}\frac{(\boldsymbol{A})_{ij}}{d_j} + \frac{1}{n^2}\sum_{j,k=1}^{n}\frac{(\boldsymbol{A})_{jk}}{d_jd_k}, \end{aligned}$$

and by applying the Cauchy–Schwarz inequality, we find $\sum_{j,k}\frac{(\boldsymbol{A})_{jk}}{d_jd_k} \leq \sum_{j,k}\frac{(\boldsymbol{A})_{jk}}{d_j^2} = \sum_j\frac{1}{d_j}$, and, therefore,

$$-\frac{2}{nh_i^{(2)}} \leq \boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i \leq \frac{1}{nh}.$$

Similarly, $\boldsymbol{e}_i'\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{M}_{\boldsymbol{\iota}}\boldsymbol{e}_i \geq 0$ contains three terms, for which we have

$$\begin{aligned} \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_i &= \frac{1}{d_ih_i}, \\ -\frac{2}{n}\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{\iota}_n &= -\frac{2}{nd_i}\sum_{j\in[i]}\frac{(\boldsymbol{A})_{ij}}{d_j}\sum_{k\in[j]}\frac{(\boldsymbol{A})_{jk}}{d_k} \leq 0, \\ \frac{1}{n^2}\boldsymbol{\iota}_n'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{\iota}_n &= \frac{1}{n^2}\sum_{i,j,k}\frac{(\boldsymbol{A})_{ij}(\boldsymbol{A})_{jk}}{d_id_jd_k} \leq \frac{1}{n^2}\sum_{i,j,k}\frac{(\boldsymbol{A})_{ij}^2}{d_i^2d_j} = \frac{1}{n}\sum_i\frac{1}{d_ih_i} = \frac{1}{hH}, \end{aligned}$$

where, in the last line, we again applied the Cauchy–Schwarz inequality, and the definitions of the harmonic means h and H in the main text. Combining the above gives the statement of Theorem S.3. *Q.E.D.*

PROOF OF THEOREM 3 (COVARIATES): Define the $n \times n$ matrix

$$C := (B'B)^{\star} B'X(X'X)^{-1} X'B.$$

Let $\lambda_i(C)$ denote the ith eigenvalue of C, arranged in ascending order. The matrix C is similar to the positive semidefinite matrix

$$(X'X)^{-1/2} X'B(B'B)^{\star} B'X(X'X)^{-1/2},$$

and since similar matrices share the same eigenvalues, we have $\lambda_1(C) \geq 0$. The matrix C is also similar to the matrix

$$B(B'B)^{\star} B'X(X'X)^{-1} X',$$

which is the product of two projection matrices, whose spectral norm is thus bounded by 1. Hence, $\lambda_n(C) \leq 1$. In addition, we must have $\lambda_i(C) \neq 1$ for any $1 < i < n$ because, otherwise, $\operatorname{rank}(I_n - C) < n$, which implies that $\operatorname{rank}(B'M_X B) < n-1$, contradicting our non-collinearity assumption (since the graph is connected, we have $\operatorname{rank}(B'B) = n-1$, which together with the non-collinearity assumption $\operatorname{rank}((X, B)) = p+n-1$ implies that $\operatorname{rank}(B'M_X B) = n-1$). We, therefore, have $\|C\|_2 < 1$, implying that $I_m - C$ is invertible.

Using (S.3) and (S.4) with $\lambda = m^{-1}$, we find that $(B'M_X B + m^{-1} D \iota_n \iota_n' D)^{-1} = (B'M_X B)^{\star} + m^{-1} \iota_n \iota_n'$ or, equivalently,

$$B'M_X B + m^{-1} D \iota_n \iota_n' D = \left[(B'M_X B)^{\star} + m^{-1} \iota_n \iota_n'\right]^{-1},$$

and analogously we have

$$B'B + m^{-1} D \iota_n \iota_n' D = \left[(B'B)^{\star} + m^{-1} \iota_n \iota_n'\right]^{-1}. \tag{S.10}$$

Subtracting the expressions in the last two displays gives

$$B'X(X'X)^{-1} X'B = \left[(B'B)^{\star} + m^{-1} \iota_n \iota_n'\right]^{-1} - \left[(B'M_X B)^{\star} + m^{-1} \iota_n \iota_n'\right]^{-1},$$

and by multiplying with $[(B'B)^{\star} + m^{-1} \iota_n \iota_n']$ from the left and $[(B'M_X B)^{\star} + m^{-1} \iota_n \iota_n']$ from the right, and using $B \iota_n = 0$, we obtain

$$(B'B)^{\star} B'X(X'X)^{-1} X'B(B'M_X B)^{\star} = (B'M_X B)^{\star} - (B'B)^{\star},$$

which can equivalently be expressed as $(I_m - C)(B'M_X B)^{\star} = (B'B)^{\star}$. We have already argued that $(I_m - C)$ is invertible and, therefore,

$$(B'M_X B)^{\star} = (I_m - C)^{-1} (B'B)^{\star}.$$

Since $\|C\|_2 < 1$, we can expand $(I_m - C)^{-1}$ in powers of C, as

$$(B'M_X B)^{\star} = \sum_{r=0}^{\infty} C^r (B'B)^{\star}. \tag{S.11}$$

Defining the $p \times p$ matrix

$$\widetilde{C} := (X'X)^{-1/2} X'B(B'B)^{\star} B'X(X'X)^{-1/2},$$

we can rewrite (S.11) as

$$(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star = (\boldsymbol{B}'\boldsymbol{B})^\star + (\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1/2}\left(\sum_{r=0}^{\infty}\widetilde{\boldsymbol{C}}^r\right)(\boldsymbol{X}'\boldsymbol{X})^{-1/2}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star.$$

The parameter ρ defined in the main text satisfies

$$\rho = \left\|(\boldsymbol{X}'\boldsymbol{X})^{-1/2}\boldsymbol{X}'\boldsymbol{M}_B\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1/2}\right\|_2 = \|\boldsymbol{I}_p - \widetilde{\boldsymbol{C}}\|_2 = 1 - \|\widetilde{\boldsymbol{C}}\|_2,$$

that is, we have $\|\widetilde{\boldsymbol{C}}\|_2 = 1 - \rho$, and since $\widetilde{\boldsymbol{C}}$ is symmetric and semidefinite, this can equivalently be written as $\widetilde{\boldsymbol{C}} \leq (1-\rho)\boldsymbol{I}_p$. Therefore,

$$\sum_{r=0}^{\infty}\widetilde{\boldsymbol{C}}^r \leq \sum_{r=0}^{\infty}(1-\rho)^r\boldsymbol{I}_p = \rho^{-1}\boldsymbol{I}_p.$$

We thus have

$$\begin{aligned}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star - (\boldsymbol{B}'\boldsymbol{B})^\star &= (\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1/2}\left(\sum_{r=0}^{\infty}\widetilde{\boldsymbol{C}}^r\right)(\boldsymbol{X}'\boldsymbol{X})^{-1/2}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star \\ &\leq \frac{1}{\rho}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star \qquad \text{(S.12)}\end{aligned}$$

and, therefore,

$$\begin{aligned}\text{var}(\check{\alpha}_i) - \text{var}(\hat{\alpha}_i) &= \sigma^2\boldsymbol{e}_i'\left[(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star - (\boldsymbol{B}'\boldsymbol{B})^\star\right]\boldsymbol{e}_i \\ &\leq \frac{\sigma^2}{\rho}\boldsymbol{e}_i'\left[(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star\right]\boldsymbol{e}_i.\end{aligned}$$

Using the expressions (S.6) and (S.7) for $(\boldsymbol{B}'\boldsymbol{B})^\star = \boldsymbol{L}^\star$, we obtain

$$\begin{aligned}&\boldsymbol{e}_i'(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{e}_i \\ &\quad = \boldsymbol{e}_i'\boldsymbol{L}^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{L}^\star\boldsymbol{e}_i \\ &\quad = \boldsymbol{e}_i'(\boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{L}^\star)\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{D}^{-1} + \boldsymbol{L}^\star\boldsymbol{A}\boldsymbol{D}^{-1})\boldsymbol{e}_i \\ &\quad \leq T_i^{(1)} + T_i^{(2)} + 2\sqrt{T_i^{(1)}T_i^{(2)}},\end{aligned}$$

where

$$\begin{aligned}T_i^{(1)} &:= \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{e}_i, \\ T_i^{(2)} &:= \boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{L}^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{L}^\star\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_i,\end{aligned}$$

and we used the Cauchy–Schwarz inequality to bound the mixed term. Again, because similar matrices have the same eigenvalues, we have

$$\left\|(\boldsymbol{L}^\star)^{1/2}\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{L}^\star)^{1/2}\right\|_2 = \|\widetilde{\boldsymbol{C}}\|_2 = 1 - \rho$$

and, therefore,

$$\begin{aligned} T_i^{(2)} &= \boldsymbol{e}_i' \boldsymbol{D}^{-1} \boldsymbol{A} (\boldsymbol{L}^\star)^{1/2} \big[(\boldsymbol{L}^\star)^{1/2} \boldsymbol{B}' \boldsymbol{X} (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{B} (\boldsymbol{L}^\star)^{1/2} \big] (\boldsymbol{L}^\star)^{1/2} \boldsymbol{A} \boldsymbol{D}^{-1} \boldsymbol{e}_i \\ &\le (1-\rho) \boldsymbol{e}_i' \boldsymbol{D}^{-1} \boldsymbol{A} \boldsymbol{L}^\star \boldsymbol{A} \boldsymbol{D}^{-1} \boldsymbol{e}_i \\ &\le \frac{1-\rho}{\lambda_2} \boldsymbol{e}_i' \boldsymbol{D}^{-1} \boldsymbol{A} \boldsymbol{D}^{-1} \boldsymbol{A} \boldsymbol{D}^{-1} \boldsymbol{e}_i \\ &= \frac{1-\rho}{\lambda_2 d_i h_i}, \end{aligned}$$

where, in the last step, we used $\boldsymbol{e}_i' \boldsymbol{D}^{-1} \boldsymbol{A} \boldsymbol{D}^{-1} \boldsymbol{A} \boldsymbol{D}^{-1} \boldsymbol{e}_i = (d_i h_i)^{-1}$. Using our definitions $\overline{\boldsymbol{x}}_i = \boldsymbol{X}' \boldsymbol{B} \boldsymbol{D}^{-1} \boldsymbol{e}_i$ and $\boldsymbol{\Omega} = \boldsymbol{X}'\boldsymbol{X}/m$, we obtain

$$T_i^{(1)} = \boldsymbol{e}_i' \boldsymbol{D}^{-1} \boldsymbol{B}' \boldsymbol{X} (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{B} \boldsymbol{D}^{-1} \boldsymbol{e}_i = \frac{1}{m} \overline{\boldsymbol{x}}_i' \boldsymbol{\Omega}^{-1} \overline{\boldsymbol{x}}_i.$$

Combining the above results, we find

$$\begin{aligned} \operatorname{var}(\check{\alpha}_i) - \operatorname{var}(\hat{\alpha}_i) &\le \frac{\sigma^2}{\rho} \Big(T_i^{(1)} + T_i^{(2)} + 2\sqrt{T_i^{(1)} T_i^{(2)}} \Big) \\ &\le \frac{\sigma^2}{\rho} \left(\frac{1}{m} \overline{\boldsymbol{x}}_i' \boldsymbol{\Omega}^{-1} \overline{\boldsymbol{x}}_i + \frac{1-\rho}{\lambda_2 d_i h_i} + 2 \sqrt{\frac{1}{m} \overline{\boldsymbol{x}}_i' \boldsymbol{\Omega}^{-1} \overline{\boldsymbol{x}}_i \frac{1-\rho}{\lambda_2 d_i h_i}} \right). \end{aligned}$$

For any $a, b \ge 0$, we have $a + b + 2\sqrt{ab} \le 2(a+b)$. Thus, a slightly cruder but simpler bound is given by

$$\big| \operatorname{var}(\check{\alpha}_i) - \operatorname{var}(\hat{\alpha}_i) \big| \le \frac{2\sigma^2}{\rho} \left(\frac{\overline{\boldsymbol{x}}_i' \boldsymbol{\Omega}^{-1} \overline{\boldsymbol{x}}_i}{m} + \frac{1-\rho}{\lambda_2 d_i h_i} \right),$$

where we also used that $\operatorname{var}(\check{\alpha}_i) \ge \operatorname{var}(\hat{\alpha}_i)$, because adding regressors can only increase the variance of the least-squares estimator under homoskedasticity. *Q.E.D.*

PROOF OF THEOREM 4 (FIRST-ORDER REPRESENTATION): Remember that we treat $\boldsymbol{B}$ and $\boldsymbol{X}$ as fixed (i.e., nonrandom) throughout. Let $\check{\boldsymbol{\beta}} := (\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{y}$. Using the model for $\boldsymbol{y}$, we find $\check{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X})^{-1} \boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{u}$. Using our assumptions $\mathbb{E}(\boldsymbol{u}) = 0$ and $\boldsymbol{\Sigma} \le \boldsymbol{I}_m \overline{\sigma}^2$, we find $\mathbb{E}(\check{\boldsymbol{\beta}} - \boldsymbol{\beta}) = 0$ and

$$\begin{aligned} \mathbb{E}\big((\check{\boldsymbol{\beta}} - \boldsymbol{\beta})(\check{\boldsymbol{\beta}} - \boldsymbol{\beta})' \big) &= \big(\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big)^{-1} \boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{\Sigma} \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big(\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big)^{-1} \\ &\le \overline{\sigma}^2 \big(\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big)^{-1} \boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{I}_m \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big(\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big)^{-1} \\ &= \overline{\sigma}^2 \big(\boldsymbol{X}' \boldsymbol{M}_{\boldsymbol{B}} \boldsymbol{X} \big)^{-1}. \end{aligned} \tag{S.13}$$

The result in (S.10) can be rewritten as

$$\boldsymbol{L}^\star = \big(\boldsymbol{L} + m^{-1} \boldsymbol{d}\boldsymbol{d}' \big)^{-1} - m^{-1} \boldsymbol{\iota}_n \boldsymbol{\iota}_n'. \tag{S.14}$$

The constrained least-squares estimator in (4) can be expressed as

$$\check{\boldsymbol{\alpha}} = \underset{\boldsymbol{a}\in\{\boldsymbol{a}\in\mathbb{R}^n:\boldsymbol{d}'\boldsymbol{a}=0\}}{\arg\min} \|\boldsymbol{y} - \boldsymbol{X}\check{\boldsymbol{\beta}} - \boldsymbol{B}\boldsymbol{a}\|^2, \tag{S.15}$$

and analogous to Theorem 1, we then find $\check{\boldsymbol{\alpha}} = \boldsymbol{L}^\star\boldsymbol{B}'(\boldsymbol{y} - \boldsymbol{X}\check{\boldsymbol{\beta}}) = (\boldsymbol{L} + m^{-1}\boldsymbol{d}\boldsymbol{d}')^{-1}\boldsymbol{B}'(\boldsymbol{y} - \boldsymbol{X}\check{\boldsymbol{\beta}})$. Multiplying by $(\boldsymbol{L} + m^{-1}\boldsymbol{d}\boldsymbol{d}')$ from the left and using our normalization $\boldsymbol{d}'\check{\boldsymbol{\alpha}} = 0$ gives

$$\boldsymbol{L}\check{\boldsymbol{\alpha}} = \boldsymbol{B}'(\boldsymbol{y} - \boldsymbol{X}\check{\boldsymbol{\beta}}).$$

Plugging $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$ and $\boldsymbol{y} = \boldsymbol{B}\boldsymbol{\alpha} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}$ into the last display, multiplying from the left with $\boldsymbol{D}^{-1}$, and rearranging terms, we obtain

$$\check{\boldsymbol{\alpha}} - \boldsymbol{\alpha} = \boldsymbol{D}^{-1}\boldsymbol{B}'\boldsymbol{u} + \boldsymbol{\epsilon} + \tilde{\boldsymbol{\epsilon}}, \tag{S.16}$$

where

$$\boldsymbol{\epsilon} := \boldsymbol{D}^{-1}\boldsymbol{A}(\check{\boldsymbol{\alpha}} - \boldsymbol{\alpha}), \qquad \tilde{\boldsymbol{\epsilon}} := -\boldsymbol{D}^{-1}\boldsymbol{B}'\boldsymbol{X}(\check{\boldsymbol{\beta}} - \boldsymbol{\beta}).$$

We have $\mathbb{E}(\check{\boldsymbol{\beta}} - \boldsymbol{\beta}) = 0$ and $\mathbb{E}(\check{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) = 0$, and, therefore, also $\mathbb{E}(\boldsymbol{\epsilon}) = \boldsymbol{0}$ and $\mathbb{E}(\tilde{\boldsymbol{\epsilon}}) = \boldsymbol{0}$. The definition $\rho = \|(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{X}\|_2$ can equivalently be written as $\rho\boldsymbol{X}'\boldsymbol{X} \geq \boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{X}$, and, therefore, $\rho^{-1}(\boldsymbol{X}'\boldsymbol{X})^{-1} \leq (\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{X})^{-1}$. Using this and (S.13), we obtain

$$\begin{aligned}\mathbb{E}(\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}') &\leq \overline{\sigma}^2\boldsymbol{D}^{-1}\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{B}}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{D}^{-1}\\ &\leq \frac{\overline{\sigma}^2}{\rho}\boldsymbol{D}^{-1}\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{D}^{-1}.\end{aligned}$$

Using $\check{\boldsymbol{\alpha}} - \boldsymbol{\alpha} = (\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{u}$ and the assumption $\boldsymbol{\Sigma} \leq \overline{\sigma}^2\boldsymbol{I}_n$, we calculate

$$\begin{aligned}\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') &= \boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{\Sigma}\boldsymbol{M}_X\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1}\\ &\leq \overline{\sigma}^2\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1}\\ &= \overline{\sigma}^2\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{M}_X\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1}\\ &\leq \overline{\sigma}^2\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1} + \frac{\overline{\sigma}^2}{\rho}\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1},\end{aligned}$$

where, in the last step, we used (S.12). Since furthermore $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \leq \boldsymbol{I}_m$ and $(\boldsymbol{B}'\boldsymbol{B})^\star = \boldsymbol{L}^\star \leq \lambda_2^{-1}\boldsymbol{D}^{-1}$, we obtain

$$\begin{aligned}\mathbb{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') &\leq \overline{\sigma}^2\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1} + \frac{\overline{\sigma}^2}{\rho}\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{B}'\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1}\\ &= \frac{\overline{\sigma}^2(1+\rho)}{\rho}\boldsymbol{D}^{-1}\boldsymbol{A}(\boldsymbol{B}'\boldsymbol{B})^\star\boldsymbol{A}\boldsymbol{D}^{-1}\\ &\leq \frac{\overline{\sigma}^2(1+\rho)}{\lambda_2\rho}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}.\end{aligned}$$

Denote the elements of $\boldsymbol{\epsilon}$ and $\tilde{\boldsymbol{\epsilon}}$ by ϵ_i and $\tilde{\epsilon}_i$. Equation (S.16) can then be written as

$$\check{\alpha}_i - \alpha_i = \frac{\boldsymbol{b}_i'\boldsymbol{u}}{d_i} + \epsilon_i + \tilde{\epsilon}_i,$$

and we have

$$\mathbb{E}(\epsilon_i^2) \leq \frac{\overline{\sigma}^2(1+\rho)}{\lambda_2\rho}\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{e}_i = \frac{\overline{\sigma}^2(1+\rho)}{\lambda_2\rho}\frac{1}{d_ih_i}$$

and

$$\mathbb{E}(\tilde{\epsilon}_i^2) \leq \frac{\overline{\sigma}^2}{\rho}\boldsymbol{e}_i'\boldsymbol{D}^{-1}\boldsymbol{B}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{e}_i = \frac{1}{m}\frac{\overline{\sigma}^2}{\rho}\overline{\boldsymbol{x}}_i'\boldsymbol{\Omega}^{-1}\overline{\boldsymbol{x}}_i,$$

where we used our definitions $\overline{\boldsymbol{x}}_i = \boldsymbol{X}'\boldsymbol{b}_i/d_i = \boldsymbol{X}'\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{e}_i$ and $\boldsymbol{\Omega} := \boldsymbol{X}'\boldsymbol{X}/m$. *Q.E.D.*

PROOF OF THEOREM 5 (ASYMPTOTIC DISTRIBUTION): We have $\rho \leq 1$ by definition. Together with the assumptions $\overline{\sigma}^2 = O(1)$, $\lambda_2 h_i \to \infty$, and the conditions in (13), this implies that $\mathbb{E}(\epsilon_i^2) \leq \overline{\sigma}^2(1+\rho)/(\rho d_i \lambda_2 h_i) = o(d_i^{-1})$ and $\mathbb{E}(\tilde{\epsilon}_i^2) \leq \overline{\sigma}^2\overline{\boldsymbol{x}}_i'\boldsymbol{\Omega}^{-1}\overline{\boldsymbol{x}}_i/(\rho m) = o(d_i^{-1})$. By Markov's inequality, we thus have $\epsilon_i = o_p(d_i^{-1/2})$ and $\tilde{\epsilon}_i = o_p(d_i^{-1/2})$, and applying Theorem 4 gives, as $d_i \to \infty$,

$$(\check{\alpha}_i - \alpha_i) \xrightarrow{p} \frac{\boldsymbol{b}_i'\boldsymbol{u}}{d_i} = \frac{1}{d_i}\sum_{j\in[i]}\sum_{e\in E_{(i,j)}} \nu_{\varepsilon_e i}, \quad \nu_{\varepsilon_e i} := (\boldsymbol{B})_{\varepsilon_e i}u_{\varepsilon_e}.$$

The number of terms $\nu_{\varepsilon_e i}$ summed over in the last display grows to infinity asymptotically, because we assume that $d_i = \sum_{j\in[i]}\sum_{e\in E_{(i,j)}} w_e \to \infty$, while the weights $w_e = (\boldsymbol{B})^2_{\varepsilon_e i}$ are bounded. Our assumptions furthermore guarantee that the $\nu_{\varepsilon_e i}$ are independent and satisfy $\mathbb{E}(\nu_{\varepsilon_e i}) = 0$, $\mathbb{E}(\nu^2_{\varepsilon_e i}) \geq c_1 > 0$ and $\mathbb{E}(|\nu_{\varepsilon_e i}|^3) \leq c_2 < \infty$ for constants c_1 and c_2. Thus, the Lyapunov condition is satisfied, and the statement of the theorem then follows from a standard application of Lyapunov's central limit theorem. *Q.E.D.*

REFERENCES

BUTLER, S. (2016): "Algebraic Aspects of the Normalized Laplacian," in *Recent Trends in Combinatorics*, ed. by A. Beveridge, J. R. Griggs, L. Hogben, G. Musiker, and P. Tetali. Springer, 295–315. [3]

CHUNG, F. R. K. (1997): *Spectral Graph Theory*. CBMS Regional Conference Series in Mathematics, Vol. 92. American Mathematical Society. [2]

FINKELSTEIN, A., M. GENTZKOW, AND H. WILLIAMS (2016): "Sources of Geographic Variation in Health Care: Evidence From Patient Migration," *Quarterly Journal of Economics*, 131, 1681–1726. [3]

Co-editor Ulrich K. Müller handled this manuscript.

Manuscript received 4 August, 2016; final version accepted 3 April, 2019; available online 10 April, 2019.