Supplementary Appendix to Inference with Many Weak Instruments

Anna Mikusheva[1] and Liyang Sun[2]

Abstract

This Supplementary Appendix contains proofs of technical results stated in the paper.

`Key words:` instrumental variables, weak identification, dimensionality asymptotics.

`JEL classification codes:` C12, C36, C55.

S1 Lemmas for sums involving projection matrix.

Lemma S1.1 *Assume that $P = (P_{ij}, i, j = 1, ..., N)$ satisfy Assumption 1, then*

(i) $|P_{ij}| \leq 1$ *and* $|M_{ij}| \leq 1$ *for any* i, j*;*

(ii) $\sum_{i'=1}^N |P_{ii'} P_{i'j}| \leq 1$ *for any* i, j*;*

(iii) $\sum_{j \neq i} P_{ij}^2 \leq \sum_{j=1}^N P_{ij}^2 = P_{ii} < 1$ *for any* i*;*

(iv) $\sum_i P_{ii}^2 \leq \sum_i P_{ii} = K$*;*

(v) $\sum_{i=1}^N P_{ii} |P_{ij}| \leq \sum_{i=1}^N \sqrt{P_{ii}} |P_{ij}| \leq \sqrt{K \cdot P_{jj}} < \sqrt{K}$ *for any* j*.*

Proof of Lemma S1.1. $M_{ij}^2 = P_{ij}^2 \leq \sum_{i'=1}^N P_{ii'}^2 = P_{ii} \leq 1$. Both M and P are non-negative definite, thus, $P_{ii} \geq 0$, thus $M_{ii} = 1 - P_{ii} \leq 1$.

$$\sum_{i'=1}^N |P_{ii'} P_{i'j}| \leq \sqrt{\sum_{i'=1}^N P_{ii'}^2} \sqrt{\sum_{i'=1}^N P_{i'j}^2} \leq \sqrt{P_{ii} P_{jj}} \leq 1,$$
$$\sum_{i=1}^N \sqrt{P_{ii}} |P_{ij}| \leq \sqrt{\sum_{i=1}^N P_{ii}} \sqrt{\sum_{i=1}^N P_{ij}^2} \leq \sqrt{K \cdot P_{jj}}.$$

[1]Department of Economics, M.I.T. Address: 77 Massachusetts Avenue, E52-526, Cambridge, MA, 02139. Email: amikushe@mit.edu. National Science Foundation support under grant number 1757199 is gratefully acknowledged.

[2]UC Berkeley and CEMFI. Email: lsun20@berkeley.edu.

Lemma S1.2 *Denote I_4 to be the set of all combinations of four indexes (i,j,i',j') where no two indexes coincide. Let Assumption 1 hold for matrix P, then:*

(a) $\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 P_{ii'}^2 P_{jj'}^2 \to 0$;

(b) $\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 |P_{ii'}P_{jj'}P_{ij'}P_{i'j}| \to 0$;

(c) $\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 |M_{ii}M_{i'i'}| P_{jj'}^2 \to 0$;

(d) $\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 |M_{ii}P_{jj'}P_{i'j'}P_{i'j}| \to 0$;

(e) $\frac{1}{K^2}\sum_{I_4} |P_{ij}^3 P_{i'j'}^3 P_{i'j}P_{ij'}| \to 0$.

Proof of Lemma S1.2. Statements (a) and (c) are proved similarly. We bound the corresponding sums by first noticing that $P_{ii'}^2 < 1$ and $|M_{ii}M_{i'i'}| < 1$, and then apply Lemma S1.1 (iii) and (iv):

$$\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 P_{jj'}^2 \le \frac{1}{K^2}\sum_{i,j,j'} P_{ij}^2 P_{j'j'} P_{jj'}^2 \le$$
$$\le \frac{1}{K^2}\sum_{i,j,j'} P_{ij}^2 P_{jj'}^2 \le \frac{1}{K^2}\sum_{i,j} P_{ij}^2 P_{jj} \le \frac{1}{K^2}\sum_{j} P_{jj}^2 \le \frac{1}{K^2}\sum_{j} P_{jj} = \frac{1}{K} \to 0.$$

Statement (b) is proved by applying Lemma S1.1 (i) and then (ii) twice:

$$\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 |P_{ii'}P_{jj'}P_{ij'}P_{i'j}| \le \frac{1}{K^2}\sum_{i,j} P_{ij}^2 \sum_{i'} |P_{ii'}P_{i'j}| \sum_{j'} |P_{jj'}P_{ij'}| \le \frac{1}{K^2}\sum_{i,j} P_{ij}^2 = \frac{1}{K} \to 0.$$

Statement (d) is proved by applying Lemma S1.1 (ii) and then (iii):

$$\frac{1}{K^2}\sum_{I_4} P_{ij}^2 P_{i'j'}^2 |M_{ii}P_{jj'}P_{i'j'}P_{i'j}| \le \frac{1}{K^2}\sum_{i',j,j'} (\sum_{i} P_{ij}^2) P_{i'j'}^2 |P_{jj'}P_{i'j}| \le \frac{1}{K^2}\sum_{i',j,j'} P_{i'j'}^2 |P_{jj'}P_{i'j}P_{jj}| \le$$
$$\le \frac{1}{K^2}\sum_{i',j'} P_{i'j'}^2 \sum_{j} |P_{jj'}P_{i'j}| \le \frac{1}{K^2}\sum_{i',j'} P_{i'j'}^2 = \frac{1}{K} \to 0.$$

Statement (e) is proved by applying Lemma S1.1 (i) and lastly (v):

$$\frac{1}{K^2}\sum_{I_4}|P_{ij}^3 P_{i'j'}^3 P_{i'j} P_{ij'}| \le \frac{1}{K^2}\sum_{i',j,j'}(\sum_i P_{ij}^2) P_{i'j'}^2 |P_{i'j}| = \frac{1}{K^2}\sum_{i',j,j'} P_{jj} P_{i'j'}^2 |P_{i'j}| =$$
$$= \frac{1}{K^2}\sum_{i',j} P_{jj} P_{i'i'} |P_{i'j}| \le \frac{1}{K^2}\sum_j P_{jj}\sqrt{\sum_{i'} P_{i'i'}^2}\sqrt{\sum_{i'} P_{i'j}^2} \le \frac{1}{K^2}\sum_j P_{jj}\sqrt{K}\cdot 1 = \frac{1}{\sqrt{K}} \to 0.$$

Lemma S1.3 *Let Assumption 1 hold for matrix P, then for any vectors a, b, c and d:*

(a) $\sum_i \sum_j P_{ij}^2 |a_i| \le \sqrt{K a'a}$;

(b) $\sum_i \sum_j P_{ij}^2 |a_i||b_j| \le \sqrt{a'ab'b}$;

(c) $\sum_i \sum_j P_{ij}^2 |a_i||b_i||c_j| \le \sqrt{a'ab'bc'c}$;

(d) $\sum_i \sum_j P_{ij}^2 |a_i||b_i||c_j||d_j| \le \sqrt{a'ab'bc'cd'd}$;

(e) $\sum_j P_{ij}^2 |a_j| \le \sqrt{P_{ii} a'a}$.

Proof of Lemma S1.3

$$\sum_i \sum_j P_{ij}^2 |a_i| \le \sum_i P_{ii}|a_i| \le \sqrt{\sum_i P_{ii}^2}\sqrt{\sum_i a_i^2} \le \sqrt{K a'a},$$
$$\sum_i \sum_j P_{ij}^2 |a_i||b_j| \le \sqrt{\sum_i \sum_j P_{ij}^2 a_i^2}\sqrt{\sum_i \sum_j P_{ij}^2 b_j^2} \le \sqrt{\sum_i P_{ii} a_i^2}\sqrt{\sum_j P_{jj} b_j^2} \le \sqrt{a'ab'b},$$
$$\sum_i \sum_j P_{ij}^2 |a_i||b_i||c_j| \le \sqrt{\sum_i \sum_j P_{ij}^2 a_i^2}\sqrt{\sum_i \sum_j P_{ij}^2 b_i^2 c_j^2} \le \sqrt{a'ab'bc'c},$$
$$\sum_i \sum_j P_{ij}^2 |a_i||b_i||c_j||d_j| \le \sqrt{\sum_i \sum_j P_{ij}^2 a_i^2 c_j^2}\sqrt{\sum_i \sum_j P_{ij}^2 d_i^2 d_j^2} \le \sqrt{a'ab'bc'cd'd},$$
$$\sum_j P_{ij}^2 |a_j| \le \sqrt{\sum_j P_{ij}^4}\sqrt{\sum_j a_j^2} \le \sqrt{\sum_j P_{ij}^2}\sqrt{a'a} = \sqrt{P_{ii} a'a}.$$

Lemma S1.4 *Let Assumption 1 holds for matrix P. Let U_i be independent random variables with $\mathbb{E}[U_i^2] < C$. Define $w_i = \sum_{j \ne i} P_{ij}\Pi_j$, where $\Pi = (\Pi_i)$ is a $N \times 1$ non-random vector. Then we have*

(a) $\max_i |w_i|^2 \le \Pi'\Pi$, $\sum_i w_i^2 \le 4\Pi'\Pi$, *and* $\sum_i w_i^4 \le 4(\Pi'\Pi)^2$;

(b) If $\frac{\Pi'\Pi}{K} \to 0$ *as* $N \to \infty$*, then* $\frac{1}{K}\sum_i w_i^2 U_i \to^p 0$.

Proof of Lemma S1.4. By the Cauchy-Schwarz inequality and Lemma S1.1:

$$\begin{aligned}
|w_i|^2 &\leq \sum_j P_{ij}^2 \sum_j \Pi_j^2 \leq P_{ii}\Pi'\Pi \leq \Pi'\Pi, \\
w_i^2 &= (P_i\Pi - P_{ii}\Pi_i)^2 \leq 2(P_i\Pi)^2 + 2P_{ii}^2\Pi_i^2, \\
\sum_i w_i^2 &\leq 2\Pi'P^2\Pi + 2\sum_i P_{ii}^2\Pi_i^2 \leq 4\Pi'\Pi, \\
\sum_i w_i^4 &\leq \max_i |w_i|^2 \sum_i w_i^2 \leq 4(\Pi'\Pi)^2, \\
\mathbb{E}\left(\frac{1}{K}\sum_i w_i^2 U_i\right)^2 &\leq \frac{C}{K^2}\sum_i w_i^4 \leq \frac{C\Pi'\Pi}{K^2} \to 0.
\end{aligned}$$

S2 Proof for consistency of the variance estimator

Lemma S2.1 *Let assumptions of Lemma 3 hold, then* $\Delta^2 A_2 \to^p 0$*, where*

$$\begin{aligned}
A_2 = &\frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i\lambda_j\xi_i\xi_j + \frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i\xi_i\Pi_j M_j\xi + \\
&+ \frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i\Pi_i\xi_j M_j\xi + \frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\Pi_i\Pi_j M_i\xi M_j\xi.
\end{aligned}$$

Proof of Lemma S2.1. Notice that the first term is mean zero, and the three last sums have non-trivial means:

$$\mathbb{E}[A_2] = \frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\left(\lambda_i\Pi_j M_{ij}\sigma_i^2 + \lambda_i\Pi_i M_{jj}\sigma_j^2 + \Pi_i\Pi_j\sum_k M_{ik}M_{jk}\sigma_k^2\right),$$

where we denote $\sigma_i^2 = \mathbb{E}\xi_i^2$. These means are negligible asymptotically:

$$\Delta^2\left|\mathbb{E}A_2\right| \leq \frac{C\Delta^2}{K}\sum_i\sum_{j\neq i}P_{ij}^2\left(|\lambda_i||\Pi_j| + |\lambda_i||\Pi_i| + |\Pi_i||\Pi_j|\right) \leq \frac{C\Delta^2\Pi'\Pi}{K} \to 0.$$

Here we apply Assumption 2, Lemma S1.1 and Lemma S1.3 (b). Consider the variance of each sum in A_2. Due to Assumption 2, the variance of the first sum in $\Delta^2 A_2$ is:

$$Var\left(\frac{\Delta^2}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i\lambda_j\xi_i\xi_j\right)\leq\frac{\Delta^4}{K^2}\sum_{i,j}P_{ij}^4\lambda_i^2\lambda_j^2\leq\frac{\Delta^4\lambda'\lambda}{K^2}\to 0.$$

The second sum in $\Delta^2 A_2$ is $\frac{\Delta^2}{K}\sum_{i,k}\left(\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i\Pi_j M_{jk}\right)\xi_i\xi_k$. It has correlated summands whenever the set of indexes (i,k) coincides. Thus the variance of this sum is bounded by

$$\frac{C\Delta^4}{K^2}\sum_{i,k}\left(\sum_j P_{ij}^2|\lambda_i\Pi_j M_{jk}|\right)^2+\frac{C\Delta^4}{K^2}\sum_{i,k}\left(\sum_j P_{ij}^2|\lambda_i\Pi_j M_{jk}|\right)\left(\sum_{j'}P_{kj'}^2|\lambda_k\Pi_{j'}M_{j'i}|\right)\leq$$
$$\leq\frac{C\Delta^4}{K^2}\left(\sum_{i,j,j'}\sum_k P_{ij}^2P_{ij'}^2|\Pi_j\Pi_{j'}\lambda_i^2||M_{jk}M_{j'k}|+\sum_{j,j'}|\Pi_j||\Pi_{j'}|\sum_{i,k}P_{ij}^2P_{kj'}^2|\lambda_i||\lambda_k|\right)\leq$$
$$\leq\frac{C\Delta^4}{K^2}\left\{\sum_i\lambda_i^2\left(\sum_j P_{ij}^2|\Pi_j|\right)^2+\sum_{j,j'}|\Pi_j||\Pi_{j'}|P_{jj}P_{j'j'}\lambda'\lambda\right\}\leq$$
$$\leq\frac{C\Delta^4}{K^2}\left(\sum_i\lambda_i^2\left(\sum_j P_{ij}^2\right)\Pi'\Pi+\lambda'\lambda\left(\sum_j P_{jj}|\Pi_j|\right)^2\right)\leq\frac{C\Delta^4}{K^2}K\Pi'\Pi\lambda'\lambda\to 0.$$

Here we apply Lemma S1.1 (ii) and the Cauchy-Schwarz inequality multiple times. The third sum in $\Delta^2 A_2$ is $\frac{\Delta^2}{K}\sum_{j,k}\sum_{i\neq j}\widetilde{P}_{ij}^2\lambda_i\Pi_i M_{jk}\xi_j\xi_k$. Its variance is bounded by

$$\frac{C\Delta^4}{K^2}\sum_{j,k}\left(\sum_i P_{ij}^2|\lambda_i\Pi_i M_{jk}|\right)^2+\left(\sum_i P_{ij}^2|\lambda_i\Pi_i M_{jk}|\right)\left(\sum_i P_{ik}^2|\lambda_i\Pi_i M_{jk}|\right)\leq$$
$$\leq\frac{C\Delta^4}{K^2}\sum_{j,k}\left(\sum_i|\lambda_i\Pi_i M_{jk}|\right)^2\leq\frac{C\Delta^4}{K^2}\sum_{j,k}M_{jk}^2\left(\sum_i|\lambda_i\Pi_i|\right)^2\leq\frac{C\Delta^4}{K^2}K\Pi'\Pi\lambda'\lambda\to 0.$$

The last sum in $\Delta^2 A_2$ is $\frac{\Delta^2}{K}\sum_{k,l}\left(\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\Pi_i\Pi_j M_{ik}M_{jl}\right)\xi_k\xi_l$. Its variance has bound

$$\frac{C\Delta^4}{K^2}\sum_{k,l}\left(\sum_{i,j}P_{ij}^2|\Pi_i\Pi_j|(|M_{ik}M_{jl}|+|M_{jk}M_{il}|)\right)^2\leq$$
$$\leq\frac{C\Delta^4}{K^2}\sum_{i,j,i',j'}P_{ij}^2P_{i'j'}^2|\Pi_i\Pi_j\Pi_{i'}\Pi_{j'}|\sum_{k,l}|M_{ik}M_{i'k}||M_{jl}M_{j'l}|\leq$$

$$\leq \frac{C\Delta^4}{K^2}\left(\sum_{ij} P_{ij}^2|\Pi_i\Pi_j|\right)^2 \leq \frac{C\Delta^4}{K^2}(\Pi'\Pi)^2 \to 0.$$

Lemma S2.2 *Let assumptions of Lemma 3 hold, then $\Delta A_1 \to^p 0$, where*

$$A_1 = \frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i\xi_i M_j\xi\xi_j + \frac{1}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\Pi_i M_i\xi\xi_j M_j\xi.$$

Proof of Lemma S2.2. A_1 has a non-trivial mean: $\mathbb{E}A_1 = \frac{4}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\Pi_i M_{ij}M_{jj}\mathbb{E}[\xi_j^3]$. Applying Lemma S1.3 (a), we note this mean vanishes under the assumptions of Lemma 3 from the paper:

$$|\Delta\mathbb{E}A_1| \leq \frac{C|\Delta|}{K}\sum_i\sum_{j\neq i}P_{ij}^2|\Pi_i| \leq \frac{C|\Delta|\sqrt{\Pi'\Pi}}{\sqrt{K}} \to 0.$$

Next, we re-write the demeaned expression as seven distinct terms:

$$\Delta(A_1 - \mathbb{E}A_1) = \frac{\Delta}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i M_{ij}\xi_i^2\xi_j + \frac{\Delta}{K}\sum_i\sum_{j\neq i}\widetilde{P}_{ij}^2\lambda_i M_{jj}\xi_i\xi_j^2 + \frac{\Delta}{K}\sum_{I_3}\widetilde{P}_{ij}^2\lambda_i M_{jk}\xi_i\xi_j\xi_k +$$
$$+\frac{\Delta}{K}\sum_{j,k}\sum_{i\neq j}\widetilde{P}_{ij}^2\Pi_i(M_{ik}M_{jj} + M_{ij}M_{jk})\xi_j^2\xi_k + \frac{\Delta}{K}\sum_j\sum_{i\neq j}\widetilde{P}_{ij}^2\Pi_i M_{ij}M_{jj}(\xi_j^3 - \mathbb{E}\xi_j^3) +$$
$$+\frac{\Delta}{K}\sum_{j,k}\left(\sum_{i\neq j}\widetilde{P}_{ij}^2\Pi_i M_{ik}M_{jk}\right)\xi_j\xi_k^2 + \frac{\Delta}{K}\sum_{(j,k,l)\in I_3}\left(\sum_{i\neq j}\widetilde{P}_{ij}^2\Pi_i M_{ik}M_{jl}\right)\xi_j\xi_k\xi_l.$$

The variances of the first two terms have the same bound (we use Lemma S1.1 (i)):

$$\frac{C\Delta^2}{K^2}\sum_{i,j}P_{ij}^4\left(\lambda_i^2 + |\lambda_i||\lambda_j|\right) \leq \frac{C\Delta^2}{K^2}\left(\sum_i(\sum_j P_{ij}^2)\lambda_i^2 + \sum_{ij}P_{ij}^2|\lambda_i||\lambda_j|\right) \leq \frac{C}{K^2}\lambda'\lambda \to 0.$$

For the third term, we notice that the two summands with indexes (i,j,k) and (i',j',k') are correlated iff $\{i,j,k\} = \{i',j',k'\}$. There are six permutations of the three indexes, for all of them except those with $\{i,j\} = \{i',j'\}$ we use Lemma S1.1 (i) to drop terms containing elements of matrix M.The variance of the third term is bounded by

$$\frac{C\Delta^2}{K^2}\sum_{I_3}\left[P_{ij}^4(\lambda_i^2 M_{jk}^2+|\lambda_i||\lambda_j||M_{ik}M_{jk}|)+P_{ij}^2P_{ik}^2(\lambda_i^2+|\lambda_i||\lambda_k|)+P_{ij}^2P_{jk}^2(|\lambda_i||\lambda_j|+|\lambda_i||\lambda_k|)\right]\leq$$

$$\leq\frac{C\Delta^2}{K^2}\left\{\sum_{i,j}P_{ij}^4(\lambda_i^2+|\lambda_i||\lambda_j|)+\sum_{i,j}P_{ij}^2\lambda_i^2+\sum_{i,k}P_{ik}^2|\lambda_i||\lambda_k|+\right.$$

$$\left.+\sum_{i,j}P_{ij}^2|\lambda_i||\lambda_j|+\sum_j\left(\sum_i P_{ij}^2|\lambda_i|\right)^2\right\}\leq\frac{C\Delta^2}{K^2}K\lambda'\lambda\to 0.$$

For the last inequality we use Lemma S1.3 (a) and (e). The variance of the fourth term is bounded by

$$\frac{C\Delta^2}{K^2}\sum_{j,k}\left(\sum_{i\neq j}\widetilde{P}_{ij}^2\Pi_i(M_{ik}M_{jj}+M_{ij}M_{jk})\right)^2\leq$$

$$\leq\frac{C\Delta^2}{K^2}\sum_{j,k}\sum_{i,i'}P_{ij}^2P_{i'j}^2|\Pi_i\Pi_{i'}|(|M_{ik}|+|M_{jk}|)(|M_{i'k}|+|M_{jk}|)\leq$$

$$\leq\frac{C\Delta^2}{K^2}\sum_{j}\sum_{i,i'}P_{ij}^2P_{i'j}^2|\Pi_i\Pi_{i'}|=\frac{C\Delta^2}{K^2}\sum_j\left(\sum_i P_{ij}^2|\Pi_i|\right)^2\leq\frac{C\Delta^2\Pi'\Pi}{K}\to 0,$$

where in the first inequality we apply Lemma S1.1 (i) to drop terms that do not index over k such as M_{jj} and $|M_{ij}|$. In the second inequality we apply Lemma S1.1 (ii) The variance of the fifth term is bounded by

$$\frac{C\Delta^2}{K^2}\sum_j\left(\sum_i P_{ij}^2|\Pi_i|\right)^2\leq\frac{C\Delta^2\Pi'\Pi}{K}\to 0.$$

The variance of the sixth term is bounded by

$$\frac{C\Delta^2}{K^2}\sum_{j,k}\left(\sum_i P_{ij}^2|\Pi_i M_{ik}M_{jk}|\right)^2+\left(\sum_i P_{ij}^2|\Pi_i M_{ik}M_{jk}|\right)\left(\sum_i P_{ik}^2|\Pi_i M_{ij}M_{jk}|\right)\leq$$

$$\leq\frac{C\Delta^2}{K^2}\sum_{j,k}M_{jk}^2\left(\sum_i|\Pi_i|\right)^2\leq\frac{C\Delta^2\Pi'\Pi}{K}\to 0.$$

Consider the seventh term that has summation over I_3. Denote o to be a permutation over indexes (j,k,l), and summation over o is the summation over all permutations. A bound on the variance of the seventh term is:

$$\frac{C\Delta^2}{K^2}\sum_{(j,k,l)\in I_3}\sum_{o}\sum_{i,i'}P_{ij}^2P_{i'o(j)}^2|\Pi_i||\Pi_{i'}||M_{ik}M_{jl}M_{i'o(k)}M_{o(j)o(l)}|.$$

Consider those permutations for which $o(j)=j$, then the term is

$$\frac{C\Delta^2}{K^2}\sum_{j}\left(\sum_{i}P_{ij}^2|\Pi_i|\right)^2\sum_{k,l}|M_{ik}M_{jl}M_{i'o(k)}M_{jo(l)}|\leq\frac{C\Delta^2}{K^2}\sum_{j}\left(\sum_{i}P_{ij}^2|\Pi_i|\right)^2\leq\frac{C\Delta^2\Pi'\Pi}{K}\to 0.$$

In the expression above, when $o(k)=l$ the summation over k and l is bounded by 1 due to Lemma S1.1 (ii). When $o(k)=k$ the summation over k is bounded by 1 due to Lemma S1.1 (ii), and the summation over l is bounded by 1 due to Lemma S1.1 (iii). Then we use Lemma S1.3 (e). Consider those permutations for which $o(j)=k$, then the term is

$$\frac{C\Delta^2}{K^2}\sum_{j,k,l}\sum_{i,i'}P_{ij}^2P_{i'k}^2|\Pi_i||\Pi_{i'}||M_{ik}M_{jl}M_{i'o(k)}M_{ko(l)}|\leq\frac{C\Delta^2}{K^2}\left(\sum_{i,j}P_{ij}^2|\Pi_i|\right)^2\leq\frac{C\Delta^2}{K}\Pi'\Pi\to 0.$$

For either $o(l)=l$ or $o(k)=l$, we apply Lemma S1.1 (ii) to the summation over l, which is bounded by 1. Then we drop all remaining M's such as $|M_{ik}|$ as they are bounded by 1 by Lemma S1.1 (i). Finally we use Lemma S1.3 (a). Consider those permutations for which $o(j)=l$ we repeat the last argument but to the index over k. To sum up, we show that all seven terms in $\Delta(A_1-\mathbb{E}A_1)$ converge in probability to zero. $\square$

S3 Statements used in Proof of Theorem 5

Lemma S3.1 *Let errors (e_i,v_i) satisfy Assumption 2, Assumption 1 hold and Π_i be such that $\Pi'M\Pi\leq\frac{C\Pi'\Pi}{K}$ and $\frac{\Pi'\Pi}{K^{2/3}}\to 0$ as $N\to\infty$. Then the following statements hold:*

(a) $\frac{1}{K}\sum_{i=1}^N\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{e_iM_ie}{M_{ii}}+\frac{1}{K}\sum_{i=1}^N\sum_{j\neq i}\widetilde{P}_{ij}^2M_iXe_iM_jXe_j-\Psi\to^p 0,$

(b) $\frac{1}{K}\sum_{i=1}^N\left(\sum_{j\neq i}P_{ij}X_j\right)^2\left(\frac{e_iM_iX}{2M_{ii}}+\frac{X_iM_ie}{2M_{ii}}\right)+\frac{1}{K}\sum_{i=1}^N\sum_{j\neq i}\widetilde{P}_{ij}^2M_iXe_iM_jXX_j-\tau\to^p 0,$

(c) $\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{X_iM_iX}{M_{ii}}+\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}\widetilde{P}_{ij}^2M_iXX_iM_jXX_j-\Upsilon\to^p 0,$

where

$$\Psi=\frac{1}{K}\sum_{i=1}^{N}(\sum_{j\neq i}P_{ij}\Pi_j)^2\sigma_i^2+\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\gamma_i\gamma_j+\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\sigma_i^2\varsigma_j^2,$$

$$\tau=\frac{2}{K}\sum_{i=1}^{N}(\sum_{j\neq i}P_{ij}\Pi_j)^2\gamma_i+\frac{2}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\varsigma_i^2\gamma_j,$$

$$\Upsilon=\frac{4}{K}\sum_{i=1}^{N}(\sum_{j\neq i}P_{ij}\Pi_j)^2\varsigma_i^2+\frac{2}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\varsigma_i^2\varsigma_j^2.$$

Proof of Lemma S3.1. Applying Lemmas 2 and 3 to different combinations of ξ_i variables containing $X_i=\Pi_i+v_i$ and e_i gives that:

$$\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}\widetilde{P}_{ij}^2M_iXe_iM_jXe_j\to\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\gamma_i\gamma_j,$$

$$\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}\widetilde{P}_{ij}^2M_iXe_iM_jXX_j\to\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\gamma_i\varsigma_j^2,$$

$$\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}\widetilde{P}_{ij}^2M_iXX_iM_jXX_j\to\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\varsigma_i^2\varsigma_j^2.$$

Thus, all that remains to prove is the convergences of the first terms in statements (a)-(c). We use $\sum_{j\neq i}P_{ij}X_j=w_i+\sum_{j\neq i}P_{ij}v_j$, where $w_i=\sum_{j\neq i}P_{ij}\Pi_j$, $X_i=\Pi_i+v_i$, and $\frac{M_ie}{M_{ii}}=e_i-\frac{1}{M_{ii}}\sum_{j\neq i}P_{ij}e_j$. Furthermore, denote $\lambda_i=M_i\Pi$.

Consider the first term in statement (a):

$$\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{e_iM_ie}{M_{ii}}=\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2e_i\left(e_i-\frac{1}{M_{ii}}\sum_{k\neq i}P_{ik}e_k\right)=$$
$$=\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2e_i^2-\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{e_i}{M_{ii}}\sum_{k\neq i}P_{ik}e_k.$$

We apply Lemma S3.2 (a) and (b) to the above, this finishes the proof of statement (a).

Consider the first term in statement (b):

$$\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\left(\frac{e_iM_iX}{M_{ii}}+\frac{X_iM_ie}{M_{ii}}\right)=$$
$$=\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\left(e_i\left[\frac{\lambda_i}{M_{ii}}+v_i-\frac{1}{M_{ii}}\sum_{k\neq i}P_{ik}v_k\right]+(\Pi_i+v_i)\left[e_i-\frac{1}{M_{ii}}\sum_{k\neq i}P_{ik}e_k\right]\right)=$$
$$=2\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2e_iv_i-\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\left\{\frac{e_i}{M_{ii}}\sum_{k\neq i}P_{ik}v_k+\frac{v_i}{M_{ii}}\sum_{k\neq i}P_{ik}e_k\right\}+$$
$$+\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2e_i\left\{\frac{\lambda_i}{M_{ii}}+\Pi_i\right\}-\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{\Pi_i}{M_{ii}}\sum_{k\neq i}P_{ik}e_k.$$

We apply Lemma S3.2 (a)-(d) to all four terms respectively. Only the first and the last terms have non-trivial limits. The first one has limit $\frac{2}{K}\sum_{i=1}^{N}w_i^2\gamma_i+\frac{2}{K}\sum_{i,j\neq i}P_{ij}^2\gamma_i\varsigma_j^2$. The last one has the limit not showing up in the expression for τ: $-\frac{1}{K}\sum_{i=1}^{N}w_i\frac{\Pi_i}{M_{ii}}\sum_{j\neq i}P_{ij}^2\gamma_k$. However, this limit is negligible as it is bounded by $\frac{C}{K}\sum_{i=1}^{N}|w_i\Pi_i|\sum_{j\neq i}P_{ij}^2\leq\frac{\Pi'\Pi}{K}\to 0$. Finally, comparing the limit with the expression for τ, we note the difference $\frac{1}{K}\sum_{i=1}^{N}w_i^2\gamma_i\leq\frac{C}{K}\sum_i w_i^2\leq\frac{\Pi'\Pi}{K}\to 0$ vanishes by Lemma S1.4 (a). This finishes the proof of (b).

Finally, we consider the first term in statement (c):

$$\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2(\Pi_i+v_i)\left(\frac{\lambda_i}{M_{ii}}+v_i-\frac{1}{M_{ii}}\sum_{k\neq i}P_{ik}v_k\right)=$$
$$=\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2v_i^2-\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{v_i}{M_{ii}}\sum_{k\neq i}P_{ik}v_k+$$
$$+\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2v_i(\Pi_i+\frac{\lambda_i}{M_{ii}})-\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\frac{\Pi_i}{M_{ii}}\sum_{k\neq i}P_{ik}v_k+$$
$$+\frac{1}{K}\sum_{i=1}^{N}\left(\sum_{j\neq i}P_{ij}X_j\right)^2\Pi_i\frac{\lambda_i}{M_{ii}}.$$

We apply Lemma S3.2 (a)-(e) to all five terms respectively. Only the first and the fourth terms have non-trivial limits. The first term has limit $\frac{1}{K}\sum_{i=1}^{N}w_i^2\varsigma_i^2+\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\varsigma_i^2\varsigma_j^2$. The fourth term has limit $\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2\varsigma_j^2\frac{\lambda_i}{M_{ii}}\Pi_i$, which does not show up in the expression for Υ, but is negligible as it is bounded by $\frac{1}{K}\sqrt{\Pi'\Pi\lambda'\lambda}\leq\frac{\Pi'\Pi}{K}\to 0$. Finally, comparing the limit with the expression for Υ, we note the difference $\frac{3}{K}\sum_{i=1}^{N} w_i^2\varsigma_i^2 \leq \frac{C}{K}\sum_i w_i^2 \leq \frac{\Pi'\Pi}{K} \to 0$ vanishes by Lemma S1.4 (a). This finishes the proof of Lemma S3.1.

Lemma S3.2 *Suppose assumptions of Lemma S3.1 hold. Let* $w_i = \sum_{j\neq i} P_{ij}\Pi_j$. *Let random variables* $\xi_{1,i}, \xi_{2,i}$ *stay for either* e_i *or* v_i, *random variables* U_i *stay for* $e_i^2, e_i v_i$ *or* v_i^2, *and constants* a_i *stay for either* Π_i *or* $\frac{\lambda_i}{M_{ii}}$. *Then the following statements hold:*

(a) $\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i} P_{ij}v_j\right)^2 U_i - \left(\frac{1}{K}\sum_{i=1}^{N} w_i^2\mathbb{E}[U_i] + \frac{1}{K}\sum_{i,j\neq i} P_{ij}^2\mathbb{E}[U_i]\varsigma_j^2\right) \to^p 0,$

(b) $\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i} P_{ij}v_j\right)^2 \frac{\xi_{1,i}}{M_{ii}}\sum_{k\neq i} P_{ik}\xi_{2,k} \to^p 0,$

(c) $\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i} P_{ij}v_j\right)^2 a_i\xi_{1,i} \to^p 0,$

(d) $\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i} P_{ij}v_j\right)^2 \frac{a_i}{M_{ii}}\sum_{k\neq i} P_{ik}\xi_{1,k} - \frac{2}{K}\sum_{i=1}^{N}\sum_{j\neq i} P_{ij}^2 w_i \frac{a_i}{M_{ii}}\mathbb{E}[v_j\xi_{1,j}] \to^p 0,$

(e) $\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i} P_{ij}v_j\right)^2 \Pi_i\frac{\lambda_i}{M_{ii}} \to^p 0.$

Proof of Lemma S3.2 For statement (a) notice that

$$\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i} P_{ij}v_j\right)^2 U_i = \frac{1}{K}\sum_{i=1}^{N} w_i^2 U_i + \frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i} P_{ij}^2 v_j^2 U_i +$$
$$+\frac{1}{K}\sum_{I_3} P_{ij}P_{ik}U_i v_j v_k + \frac{2}{K}\sum_{i=1}^{N}\sum_{j\neq i} P_{ij}w_i v_j U_i.$$

We apply Lemma S1.4 (b) to the first term. For the second term we notice that

$$\frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i} P_{ij}^2(v_j^2 U_i - \mathbb{E}[U_i]\varsigma_j^2) = \frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i} P_{ij}^2\varsigma_j^2(U_i - \mathbb{E}U_i) + \frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i} P_{ij}^2(v_j^2 - \varsigma_j^2)(U_i - \mathbb{E}U_i).$$

The summands in both sums are uncorrelated unless indexes (i in the first and i, j in the second) coincide as sets. Thus, the variance is bounded by

$$\frac{C}{K^2}\left(\sum_{i=1}^{N}\left(\sum_{j\neq i} P_{ij}^2\right)^2 + \sum_{i=1}^{N}\sum_{j\neq i} P_{ij}^4\right) \leq \frac{C}{K} \to 0.$$

The third term is

$$\frac{1}{K}\sum_{I_3} P_{ij}P_{ik}U_i v_j v_k = \frac{1}{K}\sum_{I_3} P_{ij}P_{ik}\mathbb{E}[U_i]v_j v_k + \frac{1}{K}\sum_{I_3} P_{ij}P_{ik}(U_i - \mathbb{E}U_i)v_j v_k.$$

Again the summands in both sums are uncorrelated unless indexes coincide. Thus, the variance is bounded by

$$\frac{C}{K^2}\left(\sum_{j,k}\left(\sum_i P_{ij}P_{ik}\mathbb{E}[U_i]\right)^2 + \sum_{I_3}(P_{ij}^2 P_{ik}^2 + P_{ij}^2|P_{ik}P_{jk}|)\right) \leq$$
$$\leq \frac{C}{K^2}\left(\sum_{i,i'}\left(\sum_{j,k} P_{ij}P_{ik}P_{i'j}P_{i'k}\right)\mathbb{E}[U_i]\mathbb{E}[U_{i'}] + \sum_i P_{ii}^2 + \sum_{i,j} P_{ij}^2\right) \leq \frac{C}{K} \to 0.$$

The last term is negligible as it has zero mean, and by Lemma S1.3 (b) and Lemma S1.4 (a), its variance is bounded by $\frac{C}{K}\sum_{i=1}^N \sum_{j\neq i} P_{ij}^2(w_i^2 + |w_i||w_j|) \leq \frac{Cw'w}{K} \leq \frac{C\Pi'\Pi}{K} \to 0.$

To prove statement (b) notice that the expression expands to:

$$\frac{1}{K}\sum_{i=1}^N\sum_{k\neq i} P_{ik}w_i^2\frac{\xi_{1,i}}{M_{ii}}\xi_{2,k} + \frac{2}{K}\sum_{i=1}^N\sum_{j\neq i} P_{ij}^2 w_i v_j \frac{\xi_{1,i}}{M_{ii}}\xi_{2,j} +$$
$$+\frac{2}{K}\sum_{I_3} P_{ij}P_{ik}w_i v_j\frac{\xi_{1,i}}{M_{ii}}\xi_{2,k} + \frac{1}{K}\sum_{i=1}^N\left(\sum_{j\neq i} P_{ij}v_j\right)^2\frac{\xi_{1,i}}{M_{ii}}\sum_{k\neq i} P_{ik}\xi_{2,k}.$$

All terms are mean zero. The variances of the first two are bounded by:

$$\frac{C}{K^2}\sum_{i=1}^N\sum_{k\neq i}\left(P_{ik}^2 w_i^4 + P_{ik}^2 w_i^2 w_k^2 + P_{ik}^4 w_i^2 + P_{ik}^4|w_i w_k|\right) \leq$$
$$\leq \frac{C}{K^2}\left(\max_i w_i^2 w'w + w'w\right) \leq \frac{C(\Pi'\Pi)^2}{K^2} \to 0.$$

Above we applied Lemma S1.3 (b). The variance of the third term is bounded by

$$\frac{C}{K^2}\sum_{I_3}\left(P_{ij}^2 P_{ik}^2 w_i^2 + |P_{ij}P_{ik}w_i P_{ij}P_{jk}w_j|\right) \leq \frac{C}{K^2}\left(\sum_i w_i^2 + \sum_{i,j} P_{ij}^2|w_i w_j|\right) \leq \frac{Cw'w}{K^2} \to 0.$$

Here we used Lemma S1.4 (a) and Lemma S1.3 (b).

The fourth term contains summation over i as well as summations over j, k, l where these three indexes are different from i and appear as indexes in the random variables v_j, v_l and $\xi_{2,k}$. We re-write this term as sums when all three indexes j, k, l coincide, when two of them coincide, and when all three are different. When all three indexes j, k, l coincide, the variance of that sum is bounded by $\frac{C}{K^2}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^6 \leq \frac{C}{K}$. When two of indexes j, k, l coincide (call the two distinct indexes as j, k), the variance of that sum is bounded by $\frac{C}{K^2}\sum_{I_3}\sum_{o}P_{ij}^2|P_{ik}|P_{o(i)o(j)}^2|P_{o(i)o(k)}|$ where the summation over o is the summation over all permutations of i, j, k. Consider those permutations for which $o(i) = i$, then the term is bounded by $\frac{C}{K^2}\sum_{I_3}P_{ij}^2P_{ik}^2 \to 0$. Consider those permutations for which $o(i) \neq i$, then the term is bounded by $\frac{C}{K^2}\sum_{I_3}P_{ij}^2|P_{ik}||P_{jk}| \leq \frac{C}{K^2}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2 \to 0$. Finally, when all three indexes j, k, l are distinct, the variance of that sum is bounded by $\frac{C}{K^2}\sum_{I_4}\sum_{o}|P_{ij}||P_{ik}||P_{il}||P_{o(i)o(j)}||P_{o(i)o(k)}||P_{o(i)o(l)}|$ where the summation over o is the summation over all permutations. Consider those permutations for which $o(i) = i$, then the term is bounded by $\frac{C}{K^2}\sum_{I_4}P_{ij}^2P_{ik}^2P_{il}^2 \to 0$. Consider those permutations for which $o(i) \neq i$, then it is bounded by $\frac{C}{K^2}\sum_{I_4}P_{ij}^2|P_{ik}||P_{jk}||P_{il}||P_{jl}| \leq \frac{C}{K^2}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2 \to 0$.

For proof of statement (c) we re-write this mean-zero term:

$$\frac{1}{K}\sum_{i=1}^{N}\left(w_i + \sum_{j\neq i}P_{ij}v_j\right)^2 a_i\xi_{1,i} = \frac{1}{K}\sum_{i=1}^{N}w_i^2a_i\xi_{1,i} + \frac{1}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2v_j^2a_i\xi_{1,i}+$$

$$+\frac{1}{K}\sum_{I_3}P_{ij}P_{ik}a_i\xi_{1,i}v_jv_k + \frac{2}{K}\sum_{i=1}^{N}\sum_{j\neq i}P_{ij}w_iv_ja_i\xi_{1,i}.$$

The variance of the third sum is bounded by

$$\frac{C}{K^2}\sum_{I_3}\left(P_{ij}^2P_{ik}^2a_i^2 + P_{ij}^2|P_{ik}P_{jk}a_ia_j|\right) \leq \frac{C}{K^2}\left(\sum_{i}P_{ii}^2a_i^2 + \sum_{i,j}P_{ij}^2|a_ia_j|\right) \leq \frac{Ca'a}{K^2} \to 0.$$

The variance of the remaining three terms is bounded by

$$\frac{C}{K^2}\left\{\sum_{i=1}^{N}w_i^4a_i^2 + \sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^4(a_i^2 + |a_i||a_j|) + \sum_{i=1}^{N}\sum_{j\neq i}P_{ij}^2(w_i^2a_i^2 + |w_ia_i||w_ja_j|)\right\} \leq$$

$$\leq \frac{C}{K^2}\left((\Pi'\Pi)^2a'a + a'a + (\Pi'\Pi)a'a\right).$$

We used Lemma S1.4 to derive the bound by setting a_i equal to either Π_i or $\frac{\lambda_i}{M_{ii}}$. In either case the last variance is bounded by $\frac{C(\Pi'\Pi)^3}{K^2} \to 0$.

For proof of statement (d) we expand the expression of interest to:

$$\frac{1}{K}\sum_{j=1}^{N}\left(\sum_{i\neq j} w_i^2 \frac{a_i}{M_{ii}} P_{ij}\right)\xi_{1,j} + \frac{2}{K}\sum_{j=1}^{N}\left(\sum_{i\neq j} P_{ij}^2 w_i \frac{a_i}{M_{ii}}\right)(v_j\xi_{1,j} - \mathbb{E}[v_j\xi_{1,j}]) +$$
$$+\frac{2}{K}\sum_{j=1}^{N}\sum_{k\neq j}\left(\sum_{i\notin\{k,j\}} P_{ij}P_{ik}w_i\frac{a_i}{M_{ii}}\right)v_j\xi_{1,k} + \frac{1}{K}\sum_{i=1}^{N}\frac{a_i}{M_{ii}}\left(\sum_{j\neq i}P_{ij}v_j\right)^2\sum_{k\neq i}P_{ik}\xi_{1,k}.$$

The first three terms are mean zero. The variances of the first two are bounded by

$$\frac{C}{K^2}\sum_{j=1}^{N}\sum_{i,i'}\left(w_i^2 w_{i'}^2|a_i a_{i'}P_{ij}P_{i'j}| + P_{ij}^2 P_{i'j}^2 |w_i w_{i'} a_i a_{i'}|\right) \leq$$
$$\leq \frac{C}{K^2}\left(\left(\sum_i w_i^2|a_i|\right)^2 + \left(\sum_i |w_i a_i|\right)^2\right) \leq \frac{C(\Pi'\Pi)^3}{K^2} \to 0.$$

Above we first summed up over j using Lemma S1.1 (i) and (ii), then Lemma S1.3 and finally the definition of a_i. Variance of the third term is bounded by

$$\frac{C}{K^2}\sum_{j,k}\sum_{i,i'}|P_{ij}P_{ik}w_i a_i P_{i'j}P_{i'k}w_{i'}a_{i'}| \leq \frac{C}{K^2}\left(\sum_i |w_i a_i|\right)^2 \leq \frac{C(\Pi'\Pi)^2}{K^2} \to 0.$$

The fourth term has mean $\frac{1}{K}\sum_i \sum_{j\neq i} P_{ij}^3 \frac{a_i}{M_{ii}}\mathbb{E}[v_j^2\xi_{1,j}]$, which is bounded by

$$\frac{C}{K}\sum_i P_{ii}|a_i| \leq \frac{C}{K}\sqrt{Ka'a} \leq C\sqrt{\frac{\Pi'\Pi}{K}} \to 0.$$

The de-meaned fourth term contains summation over i as well as summations over j, j', k where these three indexes appear as indexes in the random variables $v_j, v_{j'}$ and $\xi_{1,k}$. we re-write this de-meaned term as sums when all three indexes j, j', k coincide, when two of them coincide and when they all three are different. The sum of variances of these three terms are bounded by

$$\frac{C}{K^2}\left\{\sum_j\left(\sum_i |P_{ij}|^3|a_i|\right)^2+\sum_{j,k}\left(\sum_i |a_iP_{ik}|P_{ij}^2\right)^2+\sum_{j,j',k\in I_3}\left(\sum_i |a_iP_{ij}P_{ij'}P_{ik}|\right)^2\right\}.$$

For all three sums we derive the bound as follows: we write the square of the sum over i as the product of a sum over i and a sum over i', change the order of summation (moving the summation over i and i' outside). We then apply Lemma S1.1 (ii) to the summation over j, or (j,k) or I_3. Then we conclude that the expression above is bounded by

$$\frac{C}{K^2}\left(\sum_i P_{ii}|a_i|\right)^2\leq\frac{C}{K^2}\sum_i P_{ii}^2a'a\leq\frac{C\Pi'\Pi}{K}\to 0.$$

For proof of statement (e) notice:

$$\frac{1}{K}\sum_{i=1}^N\left(w_i+\sum_{j\neq i}P_{ij}v_j\right)^2\Pi_i\frac{\lambda_i}{M_{ii}}=\frac{1}{K}\sum_{i=1}^N w_i^2\Pi_i\frac{\lambda_i}{M_{ii}}+\frac{2}{K}\sum_j\left(\sum_{i\neq j}P_{ij}w_i\Pi_i\frac{\lambda_i}{M_{ii}}\right)v_j+$$
$$+\frac{1}{K}\sum_j\left(\sum_{i\neq j}P_{ij}^2\Pi_i\frac{\lambda_i}{M_{ii}}\right)v_j^2+\frac{1}{K}\sum_j\sum_{k\neq j}\left(\sum_{i\neq j,k}P_{ij}P_{ik}\Pi_i\frac{\lambda_i}{M_{ii}}\right)v_jv_k.$$

The first term is deterministic and negligible:

$$\left|\frac{1}{K}\sum_{i=1}^N w_i^2\Pi_i\frac{\lambda_i}{M_{ii}}\right|\leq\frac{C}{K}\max_i w_i^2\sqrt{\Pi'\Pi\lambda'\lambda}\leq\frac{C(\Pi'\Pi)^{3/2}(\lambda'\lambda)^{1/2}}{K}\leq\frac{C(\Pi'\Pi)^2}{K^{3/2}}\to 0.$$

The variances of the second and third term are bounded in similar fashion:

$$\frac{C}{K^2}\sum_j\left(\sum_{i\neq j}P_{ij}w_i\Pi_i\frac{\lambda_i}{M_{ii}}\right)^2\leq\frac{C}{K^2}\sum_{i,i'}\left(\sum_j|P_{ij}P_{i'j}|\right)|w_iw_{i'}\Pi_i\Pi_{i'}\lambda_i\lambda_{i'}|\leq$$
$$\leq\max_i w_i^2\frac{C}{K^2}\sum_{i,i'}\Pi_i^2\lambda_{i'}^2\leq\frac{C(\Pi'\Pi)^2\lambda'\lambda}{K^2}\leq\frac{C(\Pi'\Pi)^3}{K^3}\to 0,$$

$$\frac{C}{K^2}\sum_j\left(\sum_{i\neq j}P_{ij}^2\Pi_i\frac{\lambda_i}{M_{ii}}\right)^2\leq\frac{C}{K^2}\sum_{i,i'}(\sum_j P_{ij}^2P_{i'j}^2)|\Pi_i\Pi_{i'}\lambda_i\lambda_{i'}|\leq\frac{C\Pi'\Pi\lambda'\lambda}{K^2}\to 0.$$

Thus, the second term is negligible, while the third term converges to its mean, which happens to be negligible and is bounded by: $\frac{C}{K}\sum_j\sum_{i\neq j}P_{ij}^2\left|\Pi_i\frac{\lambda_i}{M_{ii}}\right|\leq\frac{C}{K}\sqrt{\Pi'\Pi\lambda'\lambda}\to 0$. Finally, the last term is mean zero with variance bounded by:

$$\frac{C}{K^2}\sum_j\sum_{k\neq j}\left(\sum_{i\neq j,k}P_{ij}P_{ik}\Pi_i\frac{\lambda_i}{M_{ii}}\right)^2\leq\frac{C}{K^2}\sum_{i,i'}\left(\sum_{j,k}|P_{ij}P_{ik}P_{i'j}P_{i'k}|\right)|\Pi_i\Pi_{i'}\lambda_i\lambda_{i'}|\leq\frac{C\Pi'\Pi\lambda'\lambda}{K^2}\to 0.$$

S4 Quadratic CLT for small K

Lemma S4.1 *Assume K is fixed, errors η_i are independently drawn with $\mathbb{E}[\eta_i]=0,\mathbb{E}[\eta_i^2]=\sigma^2$ and $\max_i\mathbb{E}\eta_i^4<C$. Assume also that as $N\to\infty$ the $K\times 1$-dimensional instruments Z_i satisfy the following convergence $\frac{1}{N}\sum_{i=1}^N Z_iZ_i'\to Q$, where Q is a full rank $K\times K$ matrix, and $\frac{1}{N}\sum_{i=1}^N\|Z_i\|^4<C$. Then as $N\to\infty$*

$$\frac{1}{\sqrt{K}\sqrt{\Phi}}\sum_{i=1}^N\sum_{j\neq i}P_{ij}\eta_i\eta_j\Rightarrow\frac{\chi_K^2-K}{\sqrt{2K}}.$$

Proof of Lemma S4.1. Under homoscedasticity we have $\Phi_N=2\sigma^4\cdot(1-\frac{\sum_{i=1}^N P_{ii}^2}{K})$, but we show later $\sum_{i=1}^N P_{ii}^2\to 0$, thus $\Phi=2\sigma^4$. Below we use $\sum_{i=1}^N P_{ii}=K$.

$$\frac{1}{\sqrt{2K}\sigma^2}\sum_{i=1}^N\sum_{j\neq i}P_{ij}\eta_i\eta_j=\frac{1}{\sqrt{2K}\sigma^2}\left\{\eta'Z(Z'Z)^{-1}Z'\eta-K\sigma^2\right\}-\frac{1}{\sqrt{2K}}\sum_{i=1}^N P_{ii}\left(\frac{\eta_i^2}{\sigma^2}-1\right).$$

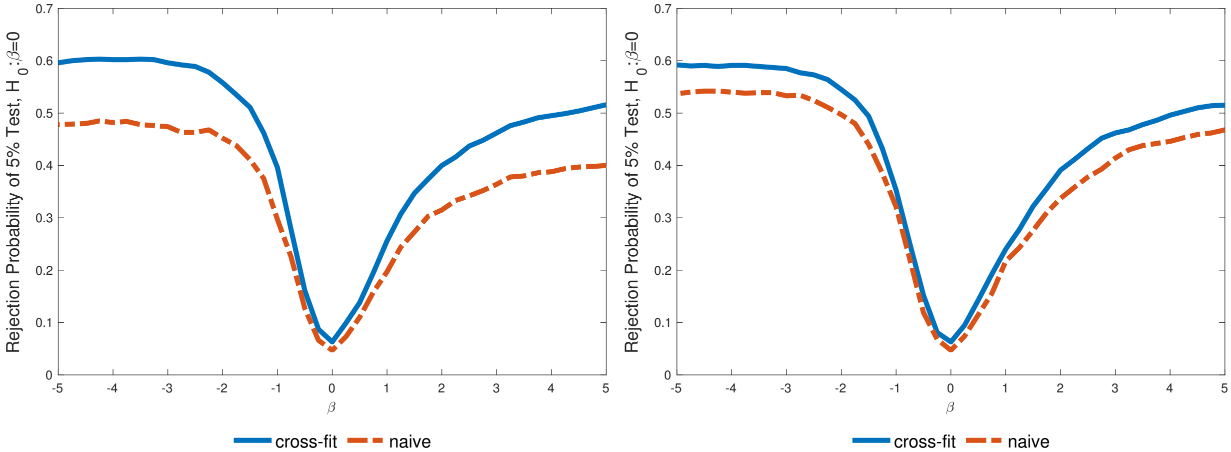
By the standard argument we have $\frac{1}{\sqrt{N}}Z'\eta\Rightarrow N(0,\sigma^2Q)$, and thus,

$$\frac{1}{\sigma^2}\eta'Z(Z'Z)^{-1}Z'\eta\Rightarrow\chi_K^2.$$

Noticing that $\frac{1}{N}\sum_{i=1}^N Z_iZ_i'\to Q$, where Q is a full rank, we have

$$\begin{aligned}P_{ii}=&Z_i'(Z'Z)^{-1}Z_i\leq\frac{\|Z_i\|^2}{N}\mathrm{tr}\left[\left(\frac{Z'Z}{N}\right)^{-1}\right]\leq\frac{C\|Z_i\|^2}{N},\\ \sum_{i=1}^N P_{ii}^2&\leq\frac{C}{N^2}\sum_{i=1}^N\|Z_i\|^4\leq\frac{C}{N}\to 0.\end{aligned}$$

Thus, by Chebyshev's inequality we have $\frac{1}{\sqrt{2K}}\sum_{i=1}^N P_{ii}\left(\frac{\eta_i^2}{\sigma^2}-1\right)\to^p 0$ as $N\to\infty$. $\square$



(a) sparse, $\frac{\mu^2}{\sqrt{K}} = 2.5$ (b) dense, $\frac{\mu^2}{\sqrt{K}} = 2.5$

Figure 1: Power curves for leave-one-out AR tests with cross-fit (blue line) and naive (red dash) variance estimators under sparse vs. dense first stage. The instruments are $K = 40$ balanced group indicators. Sample size is $N = 200$. Number of simulation draws is 1,000. Details of the simulation design can be found in the Appendix.

S5 Additional Simulations

Here we report additional simulations to the ones reported in Section 4.2 about the effect of naive vs cross-fit variance estimator on the power of the AR test. We consider the following simulation design. The DGP is given by a homoscedastic linear IV model with a linear first stage:

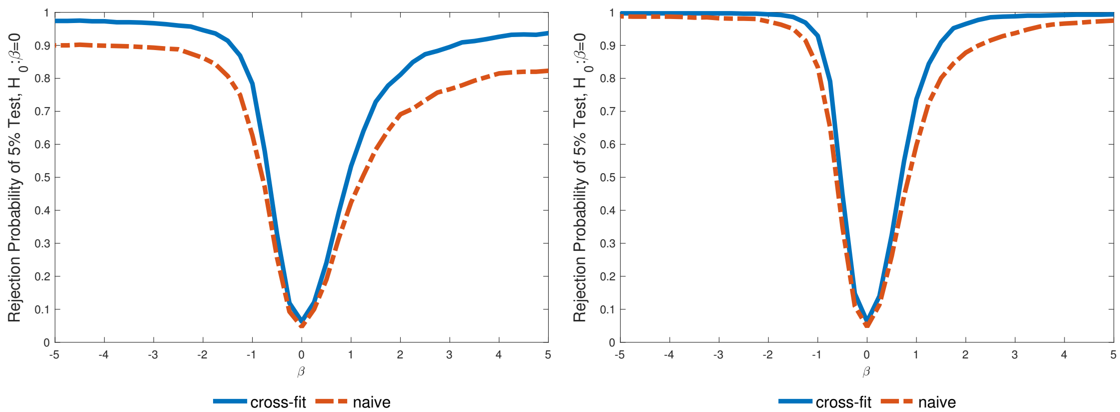
$$\begin{cases} Y_i = \beta X_i + e_i, \\ X_i = \Pi' Z_i + v_i. \end{cases}$$

The instruments are $K = 40$ group indicators, where the sample is divided into equal groups. The sample size is $N = 200$. The error terms are generated i.i.d. as

$$\begin{pmatrix} e_i \\ v_i \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

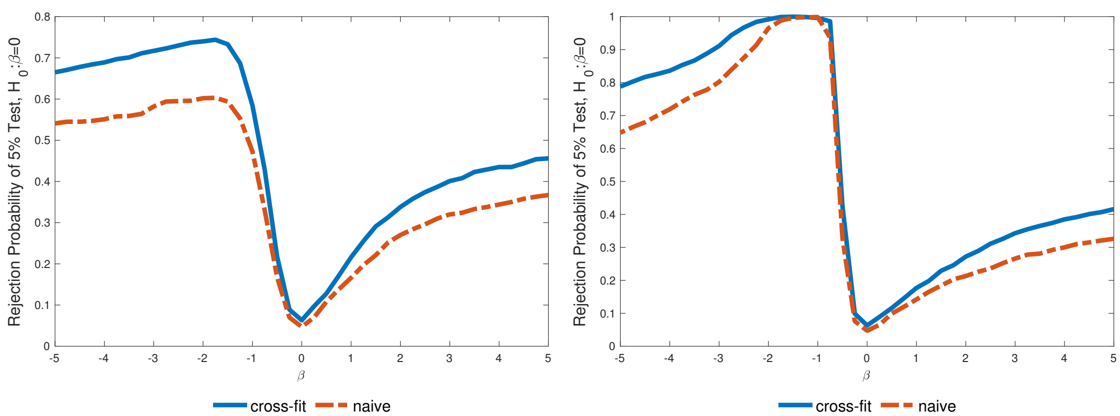
with $\rho = 0.2$. We simulate a sparse first stage by setting one large coefficient $\pi_K = 2$ and $\pi_k = 0.001$ for all $k < K$. The dense first stage has homogeneous first stage coefficients $\pi_k = 0.316$ for all $k = 1, \ldots, K$. Identification strength is held the same at $\frac{\mu^2}{\sqrt{K}} = 2.5$ for both settings. The results are reported in Figure 1.

As we discuss in the main text, the power difference between tests with the cross-fit and



(a) sparse, $\frac{\mu^2}{\sqrt{K}} = 5.7$ (b) sparse, $\frac{\mu^2}{\sqrt{K}} = 8.2$

Figure 2: Power curves for leave-one-out AR tests with cross-fit (blue line) and naive (red dash) variance estimators under sparse first stage. The instruments are $K = 40$ balanced group indicators. Sample size is $N = 200$. Number of simulation draws is 1000.



(a) sparse, $\frac{\mu^2}{\sqrt{K}} = 2.5$, $\rho = 0.5$ (b) sparse, $\frac{\mu^2}{\sqrt{K}} = 2.5$, $\rho = 0.9$

Figure 3: Power curves for leave-one-out AR tests with cross-fit (blue line) and naive (red dash) variance estimators under sparse first stage. The instruments are $K = 40$ balanced group indicators. Sample size is $N = 200$. Number of simulation draws is 1000.

the naive variance estimators is less pronounced when identification is strong. Figure 2 illustrates this by considering the same sparse design as in Figure 1, but with $\pi_K = 3$ in plot (a) and $\pi_K = 3.6$ in plot (b). These settings correspond to stronger identification as measured by $\frac{\mu^2}{\sqrt{K}}$.

Interestingly enough, the level of endogeneity changes the shape of the power curves, but not the power comparison between the two tests. Figure 3 reports results for the same sparse setting as in Figure 1, but with moderate ($\rho = 0.5$) and strong ($\rho = 0.9$) endogeneity environments.