On-Line Appendix for: "Consistency of the OLS Bootstrap for Independently but Not-Identically Distributed Data: A Permutation Perspective."

Alwyn Young, London School of Economics, July 2022

A. Multivariate extension of Theorem I

Following the presentation in the paper, let $\boldsymbol{\delta}' = (d_1, \ldots, d_N)$ denote a sequence of real numbers, $\mathbf{Z}' = (\mathbf{z}_1,\ldots,\mathbf{z}_N)$ a sequence of K x 1 real vectors, and $\mathbf{O} = \mathbf{I}_N - \mathbf{1}_N\mathbf{1}_N'/N$ the centering matrix. We wish to show that across the row permutations $\mathbf{d}$ of $\boldsymbol{\delta}$

$$\text{(A1)}\quad \mathbf{v}(\mathbf{Z},\boldsymbol{\delta}) = \left(\frac{\mathbf{Z}'\mathbf{O}\mathbf{Z}}{N}\frac{\mathbf{d}'\mathbf{O}\mathbf{d}}{N}\right)^{-\frac{1}{2}}\frac{(\mathbf{Z}'\mathbf{O}\mathbf{d})}{\sqrt{N}}$$

is asymptotically distributed multivariate iid standard normal if for all N sufficiently large $\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}'$ is non-zero and the correlation matrix $\mathbf{D}_{\mathbf{Z}'\mathbf{O}\mathbf{Z}}^{-\frac{1}{2}}\mathbf{Z}'\mathbf{O}\mathbf{Z}\mathbf{D}_{\mathbf{Z}'\mathbf{O}\mathbf{Z}}^{-\frac{1}{2}}$, where $\mathbf{D}_{\mathbf{Z}'\mathbf{O}\mathbf{Z}}$ is the diagonal matrix with diagonal entries equal to those of $\mathbf{Z}'\mathbf{O}\mathbf{Z}$, is non-singular with determinant > Δ (a positive constant), while

$$\text{(A2)}\quad \lim_{N\to\infty}\frac{N^{\frac{\tau}{2}-1}\sum_{i=1}^{N}[z_{ik}-m(z_{ik})]^{\tau}\sum_{i=1}^{N}[d_i-m(d_i)]^{\tau}}{\left(\sum_{i=1}^{N}[z_{ik}-m(z_{ik})]^2\right)^{\tau/2}\left(\sum_{i=1}^{N}[d_i-m(d_i)]^2\right)^{\tau/2}} = 0.$$

for each element z_{ik} in the vector sequence $\mathbf{z}_i$. Hoeffding (1951) provides a proof for a broader, but univariate, permutation problem. The generalization to the multivariate case requires additional notation, but otherwise I keep the presentation as close as possible to Hoeffding's so that the proof can be checked against his original contribution if desired.

Define

$$\text{(A3)}\ \ \widetilde{\mathbf{Z}} = \mathbf{OZ}\left(\frac{\mathbf{Z}'\mathbf{OZ}}{N}\right)^{-1/2} \ \ \&\ \ \widetilde{\mathbf{d}} = \mathbf{Od}\left(\frac{\mathbf{d}'\mathbf{Od}}{N}\right)^{-1/2},\ \text{so that}\ \mathbf{v}(\mathbf{Z},\boldsymbol{\delta}) = \frac{\widetilde{\mathbf{Z}}'\widetilde{\mathbf{d}}}{\sqrt{N}}.$$

For the k^{th} element of $\mathbf{v}$ we know that

$$\text{(A4)}\ \ v_k = \sum_{i=1}^{N}\frac{\widetilde{z}_{ik}\widetilde{d}_i}{N^{1/2}},\ \text{where} \sum_{i=1}^{N}\widetilde{z}_{ik} = \sum_{i=1}^{N}\widetilde{d}_i = 0\ \ \&\ \ \sum_{i=1}^{N}\widetilde{z}_{ik}^2 = \sum_{i=1}^{N}\widetilde{d}_i^2 = N\ \ \text{for all}\ k,$$

$$\text{and}\ \sum_{i=1}^{N}\widetilde{z}_{ik_1}\widetilde{z}_{ik_2} = 0\ \ \text{for all}\ k_1 \neq k_2,\ \text{as}\ \mathbf{1}'_N\widetilde{\mathbf{Z}} = \mathbf{1}'_N\widetilde{\mathbf{d}} = 0,\ \widetilde{\mathbf{Z}}'\widetilde{\mathbf{Z}} = N * \mathbf{I}_N\ \ \&\ \ \widetilde{\mathbf{d}}'\widetilde{\mathbf{d}} = N.$$

We shall show that all of the moments of the vector $\mathbf{v}$ converge to those of the mean zero multivariate normal with identity covariance matrix.

We begin by showing how the moments of the permuted variables are calculated. As $\mathbf{d}$ is the row permutation of $\boldsymbol{\delta}$, $\widetilde{\mathbf{d}} = \mathbf{Od}(\mathbf{d}'\mathbf{Od}/N)^{-1/2}$ is simply the row permutation of $\widetilde{\boldsymbol{\delta}} = \mathbf{O}\boldsymbol{\delta}(\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}/N)^{-1/2}$ and the sample moments of $\widetilde{\mathbf{d}}$ are the same as those of $\widetilde{\boldsymbol{\delta}}$. From the symmetry of the permutations, each element of $\widetilde{\mathbf{d}}$ has the same distribution, with expectations across permutations $\mathbf{d}$ given by

$$\text{(A5)}\ \ E_{\mathbf{d}}(\widetilde{d}_i) = \sum_{j=1}^{N}\frac{\widetilde{d}_j}{N} = 0\ \ \&\ \ E_{\mathbf{d}}(\widetilde{d}_i^2) = \sum_{j=1}^{N}\frac{\widetilde{d}_j^2}{N} = 1,$$

while if $i_1 \neq i_2$ we have

$$\text{(A6)}\ \ E_{\mathbf{d}}(\widetilde{d}_{i_1}\widetilde{d}_{i_2}) = \sum_{j_1,j_2=1}^{N}\frac{\widetilde{d}_{j_1}\widetilde{d}_{j_2}}{N(N-1)} = \sum_{j_1=1}^{N}\sum_{j_2=1}^{N}\frac{\widetilde{d}_{j_1}\widetilde{d}_{j_2}}{N(N-1)} - \sum_{j_1=1}^{N}\frac{\widetilde{d}_{j_1}^2}{N(N-1)} = 0 - \frac{1}{N-1},$$

where we use the notation $j_1, j_2,...$ to denote summation across multiple indices, excluding ties between the indices. Using these, we compute the 1[st] and 2[nd] moments of the components of $\mathbf{v}$:

$$\text{(A7)}\quad E_{\mathbf{d}}(v_k)=\sum_{i=1}^{N}\frac{E_{\mathbf{d}}(\widetilde{d}_i)\widetilde{z}_{ik}}{N^{½}}=\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\widetilde{d}_j\widetilde{z}_{ik}}{N^{3/2}}=0$$

$$E_{\mathbf{d}}(v_{k_1}v_{k_2})=\sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\frac{E_{\mathbf{d}}(\widetilde{d}_{i_1}\widetilde{d}_{i_2})\widetilde{z}_{i_1k_1}\widetilde{z}_{i_2k_2}}{N}=\sum_{i_1=1}^{N}\frac{E_{\mathbf{d}}(\widetilde{d}_{i_1}^2)\widetilde{z}_{i_1k_1}\widetilde{z}_{i_1k_2}}{N}+\sum_{i_1\neq i_2=1}^{N}\frac{E_{\mathbf{d}}(\widetilde{d}_{i_1}\widetilde{d}_{i_2})\widetilde{z}_{i_1k_1}\widetilde{z}_{i_2k_2}}{N}$$

$$=\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\frac{\widetilde{d}_{j_1}^2\widetilde{z}_{i_1k_1}\widetilde{z}_{i_1k_2}}{N^2}+\sum_{i_1\neq i_2=1}^{N}\sum_{j_1\neq j_2=1}^{N}\frac{\widetilde{d}_{j_1}\widetilde{d}_{j_2}\widetilde{z}_{i_1k_1}\widetilde{z}_{i_2k_2}}{N^2(N-1)}$$

$$=\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\frac{\widetilde{d}_{j_1}^2\widetilde{z}_{i_1k_1}\widetilde{z}_{i_1k_2}}{N^2}+\underbrace{\sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\sum_{j_1=1}^{N}\sum_{j_2=1}^{N}\frac{\widetilde{d}_{j_1}\widetilde{d}_{j_2}\widetilde{z}_{i_1k_1}\widetilde{z}_{i_2k_2}}{N^2(N-1)}}_{=0}$$

$$-\underbrace{\sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\sum_{j_1=1}^{N}\frac{\widetilde{d}_{j_1}^2\widetilde{z}_{i_1k_1}\widetilde{z}_{i_2k_2}}{N^2(N-1)}}_{=0}-\underbrace{\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\sum_{j_2=1}^{N}\frac{\widetilde{d}_{j_1}\widetilde{d}_{j_2}\widetilde{z}_{i_1k_1}\widetilde{z}_{i_1k_2}}{N^2(N-1)}}_{=0}+\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\frac{\widetilde{d}_{j_1}^2\widetilde{z}_{i_1k_1}\widetilde{z}_{i_1k_2}}{N^2(N-1)}$$

$$=1+\frac{1}{N-1}\ \text{(if } k_1=k_2\text{) or 0 (otherwise).}$$

These examples illustrate, in a manner that hopefully makes the later exposition intelligible, how the calculation of expectations produces sums of summations, with those that are across unequal indices in turn expressible as further sums of summations. In the more immediate sense, (A7) shows that the first moment of **v** is **0**, while its second moments asymptotically equal the identity matrix, as desired. The next few pages focus on the higher moments.

Let $E_{\mathbf{d}}^{\tau}$ denote one of the τ^{th} moments of the joint distribution of **v** across the row permutations **d** of **δ**

$$\text{(A8)}\quad E_{\mathbf{d}}^{\tau}=E_{\mathbf{d}}\left[\prod_{p=1}^{\tau}v_{k_p}\right]=E_{\mathbf{d}}\left[N^{-\tau/2}\sum_{i_1=1}^{N}\ldots\sum_{i_\tau=1}^{N}\widetilde{d}_{i_1}\widetilde{z}_{i_1k_1}\ldots\widetilde{d}_{i_\tau}\widetilde{z}_{i_\tau k_p}\right],$$

where the k_p indices may reference the same columns of $\widetilde{\mathbf{Z}}$, i.e. $k_p = k_q$ for some $p \neq q$, so that the moment is across combinations of powers of the v_k. As can be seen from the second line of (A7) earlier above, $E_{\mathbf{d}}^{\tau}$ needs to be separated into components based upon whether the i indices tie or not, which leads to elements of the form

$$\text{(A9)}\;\; I(\tau,\{e_1\},...,\{e_m\}) = E_{\mathbf{d}}\left[N^{-\tau/2}\sum_{i_1,...,i_m=1}^{N}\tilde{d}_{i_1}^{e_1}\tilde{z}_{i_1}^{\{e_1\}}...\tilde{d}_{i_m}^{e_m}\tilde{z}_{i_m}^{\{e_m\}}\right], \text{ where } \sum_{i=1}^{m} e_i = \tau,\; e_i \geq 1\,\forall\, i,$$

and $\sum_{i_1,...,i_m}$ denotes the summation across m indices, excluding ties between the indices, the sets $\{e_1\}$, ... ,$\{e_m\}$ constitute a partition of the τ v_k used in $E_{\mathbf{d}}^{\tau}$, with the notation e_i without {} denoting the number of elements in $\{e_i\}$, and the $\tilde{d}^{e_i}$ and $\tilde{z}^{\{e_i\}}$ denoting the product of the elements within each set $\{e_i\}$. The $\{e_i\}$ groupings tie elements together through their i and j indices. Thus, for example, we might have

$$\text{(A10)}\;\; \{e_1\} = \{v_{k_1}, v_{k_2}\}, \{e_2\} = \{v_{k_3}\},...,\{e_m\} = \{v_{k_{\tau-1}}, v_{k_\tau}\}$$
$$\tilde{d}_{i_1}^{e_1} = \tilde{d}_{i_1}^{2},\; \tilde{d}_{i_2}^{e_2} = \tilde{d}_{i_2},\; ...\; , \tilde{d}_{i_m}^{e_m} = \tilde{d}_{i_m}^{2}$$
$$\text{and }\; \tilde{z}_{i_1}^{\{e_1\}} = \tilde{z}_{i_1k_1}\tilde{z}_{i_1k_2},\; \tilde{z}_{i_2}^{\{e_2\}} = \tilde{z}_{i_2k_3},\; ...\; , \tilde{z}_{i_m}^{\{e_m\}} = \tilde{z}_{i_mk_{\tau-1}}\tilde{z}_{i_mk_\tau}$$

Since

$$\text{(A11)}\;\; E_{\mathbf{d}}\left[\tilde{d}_{i_1}^{e_1}...\tilde{d}_{i_m}^{e_m}\right] = \frac{N-m!}{N!}\sum_{j_1,...,j_m}^{N}\tilde{d}_{j_1}^{e_1}...\tilde{d}_{j_m}^{e_m}\;,$$

we have

$$\text{(A12)}\;\; I(\tau,\{e_1\},...,\{e_m\}) = \underbrace{\frac{N-m!N^m}{N!}}_{\to 1} N^{-m-\frac{\tau}{2}}\sum_{i_1,...,i_m}^{N}\sum_{j_1,...,j_m}^{N}\tilde{d}_{j_1}^{e_1}\tilde{z}_{i_1}^{\{e_1\}}...\tilde{d}_{j_m}^{e_m}\tilde{z}_{i_m}^{\{e_m\}}$$
$$\sim N^{-m-\frac{\tau}{2}}\sum_{i_1,...,i_m}^{N}\sum_{j_1,...,j_m}^{N}\tilde{d}_{j_1}^{e_1}\tilde{z}_{i_1}^{\{e_1\}}...\tilde{d}_{j_m}^{e_m}\tilde{z}_{i_m}^{\{e_m\}},$$

which in turn can be expressed as the sum and difference of terms of the form

$$\text{(A13)}\;\; N^{-m-\frac{\tau}{2}}J(\tau,p,q,\{e_1\},...,\{e_m\}) = N^{-m-\frac{\tau}{2}}\sum_{i_1=1}^{N}...\sum_{i_p=1}^{N}\sum_{j_1=1}^{N}...\sum_{j_q=N}^{N}\tilde{d}_{j_{d_1}}^{e_1}\tilde{z}_{i_{c_1}}^{\{e_1\}}...\tilde{d}_{j_{d_m}}^{e_m}\tilde{z}_{i_{c_m}}^{\{e_m\}}$$
$$\text{with }\; 1 \leq p \leq m,\; 1 \leq q \leq m,\; 1 \leq c_g \leq m,\; 1 \leq d_h \leq m,\; (g,h = 1,...,m)$$

and at least one c_g (d_h) equal to every integer in 1..p (1...q) . The 2m indices c_g and d_h connect the 2m different elements to the distinct $p \leq m$ and $q \leq m$ counters in the summations. The third line of (A7) earlier provides an example of how expectations add summations across j to each $I(\tau,...)$, while the fourth and fifth lines show how the $I(\tau,...)$ are re-expressed as the sum of $J(\tau,...)$ forms.

Each J can be written as the product of subset J's

$$\text{(A14)}\quad J(\tau, p, q, \{e_1\},...,\{e_m\}) = \prod_{u=1}^{s} J(\tau_u, p_u, q_u, \{e_{u1}\},...,\{e_{um_u}\})$$

where each $\{e_{ua}\}$ equals one of the original $\{e_i\}$, and the s sets $\{e_{u1}\},...,\{e_{um_u}\}$ cover $\{e_1\},...,\{e_m\}$ in its entirety, with

$$\text{(A15)}\quad \sum_{i=1}^{m_u} e_{ui} = \tau_u,\ \sum_{u=1}^{s} \tau_u = \tau,\ \sum_{u=1}^{s} p_u = p,\ \sum_{u=1}^{s} q_u = q,\ \&\ \sum_{u=1}^{s} m_u = m.$$

We assume that each J is subdivided into the greatest possible number of factors. In the fourth line of (A7) above, for example, we have:

$$\text{(A16)}\quad J(\tau = 2, p = 2, q = 2, \{v_{k_1}\}, \{v_{k_2}\}) = \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\sum_{j_1=1}^{N}\sum_{j_2=1}^{N} \tilde{d}_{j_1}\tilde{d}_{j_2}\tilde{z}_{i_1k_1}\tilde{z}_{i_2k_2} =$$

$$\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\tilde{d}_{j_1}\tilde{z}_{i_1k_1}\sum_{i_2=1}^{N}\sum_{j_2=1}^{N}\tilde{d}_{j_2}\tilde{z}_{i_2k_2} = J(\tau_1 = 1, p_1 = 1, q_1 = 1, \{v_{k_1}\})J(\tau_2 = 1, p_2 = 1, q_2 = 1, \{v_{k_2}\})$$

while all three terms in the fifth line are indivisible because the i, j counters for the $\tilde{d}$ and $\tilde{z}$ elements connect at least one element of v_{k_1} to v_{k_2}. If $J(\tau_u, p_u, q_u, \{e_{u1}\},...,\{e_{um_u}\})$ is indivisible, it is because the $2m_u$ c_{ug} and d_{uh} subscript indices link across the m_u groups $\{e_{u1}\},...,\{e_{um_u}\}$. To do so, there must be at least m_u-1 equalities in these indices, i.e. at most $m_u + 1$ distinct values. At the same time, these indices cover every one of the numbers in 1...p_u and 1...q_u, so we may conclude that

$$\text{(A17)}\quad p_u + q_u \leq m_u + 1.$$

We note that if $(c_{ug},d_{ug}) = (c_{uh},d_{uh})$ for some $ug \neq uh$, we have more than the minimum m_u-1 equalities necessary for indivisibility and (A17) holds with strict inequality. Summing across all s groups that make up $J(\tau, p, q, \{e_1\},...,\{e_m\})$,

$$\text{(A18)}\quad p + q \leq m + s$$

with strict inequality if $(c_{ug},d_{ug}) = (c_{uh},d_{uh})$ ever holds.

Next, we take the absolute value, apply an inequality associated with that, and then apply Hölder's Inequality as well:

$$\text{(A19)}\quad \left|J(\tau_u,p_u,q_u,\{e_{u1}\},...,\{e_{um_u}\})\right| \le \sum_{i_1=1}^{N}...\sum_{i_{p_u}=1}^{N}\sum_{j_1=1}^{N}...\sum_{j_{q_u}=N}^{N}\left|\tilde{d}_{j_{d_1}}^{e_{u1}}\tilde{z}_{i_{c_1}}^{\{e_{u1}\}}\right|...\left|\tilde{d}_{j_{d_{m_u}}}^{e_{um_u}}\tilde{z}_{i_{c_{m_u}}}^{\{e_{um_u}\}}\right|$$

$$\le \prod_{g=1}^{m_u}\left(\sum_{i_1=1}^{N}...\sum_{i_{p_u}=N}^{N}\sum_{j_1=1}^{N}...\sum_{j_{q_u}=N}^{N}\left|\tilde{d}_{j_{d_g}}^{e_{ug}}\tilde{z}_{i_{c_g}}^{\{e_{ug}\}}\right|^{\tau_u/e_{ug}}\right)^{e_{ug}/\tau_u} = \prod_{g=1}^{m_u}\left(N^{p_u+q_u-2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{e_{ug}}\tilde{z}_i^{\{e_{ug}\}}\right|^{\tau_u/e_{ug}}\right)^{e_{ug}/\tau_u}$$

where the reader is reminded that e_{ug} denotes the number of v_k in $\{e_{ug}\}$, with $\sum e_{ug} = \tau_u$, allowing the application of Hölder's Inequality in the manner shown. We now decompose the set $\{e_{ug}\}$ into its constituent parts. Let $1..r$, $r \le \tau$, index the unique v_k variables across which the expectation $E_{\mathbf{d}}^{\tau}$ is taken, so that

$$\text{(A20)}\quad E_{\mathbf{d}}^{\tau} = E_{\mathbf{d}}\left[\prod_{p=1}^{\tau} v_{k_p}\right] = E_{\mathbf{d}}\left[\prod_{a=1}^{r} v_{k_a}^{f_a}\right],$$

where, as earlier above, in the first product different k_p may reference the same v_k, but in the second product each a references a unique v_k, with $f_{1ug}...f_{rug}$ denoting the power the unique v_k are raised to in the grouping $\{e_{ug}\}$. We then apply Hölder's Inequality once again

$$\text{(A21)}\quad \prod_{g=1}^{m_u}\left(N^{p_u+q_u-2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{e_{ug}}\tilde{z}_i^{\{e_{ug}\}}\right|^{\tau_u/e_{ug}}\right)^{e_{ug}/\tau_u}$$

$$= N^{p_u+q_u-2}\prod_{g=1}^{m_u}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{f_{1ug}}...\tilde{d}_j^{f_{rug}}\tilde{z}_{ik_1}^{f_{1ug}}...\tilde{z}_{ik_r}^{f_{rug}}\right|^{\tau_u/e_{ug}}\right)^{e_{ug}/\tau_u} \text{ where } \sum_{g=1}^{m_u}\sum_{a=1}^{r} f_{aug} = \sum_{g=1}^{m_u} e_{ug} = \tau_u,$$

$$\le N^{p_u+q_u-2}\prod_{g=1}^{m_u}\left(\prod_{a=1}^{r}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{\tau_u}\tilde{z}_{ik_a}^{\tau_u}\right|\right)^{f_{aug}/e_{ug}}\right)^{e_{ug}/\tau_u} = N^{p_u+q_u-2}\prod_{g=1}^{m_u}\prod_{a=1}^{r}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{\tau_u}\tilde{z}_{ik_a}^{\tau_u}\right|\right)^{f_{aug}/\tau_u}$$

$$= N^{p_u+q_u-2}\prod_{a=1}^{r}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{\tau_u}\tilde{z}_{ik_a}^{\tau_u}\right|\right)^{f_{au}/\tau_u} \text{ where } f_{au} = \sum_{g=1}^{m_u} f_{aug} \;\&\; \sum_{a=1}^{r} f_{au} = \tau_u.$$

$$= N^{p_u+q_u+\frac{\tau_u}{2}-1}\prod_{a=1}^{r}\overline{M}(\tau_u,v_{k_a})^{f_{au}/\tau_u} \text{ where } \overline{M}(\tau_u,v_{k_a}) = N^{-\frac{\tau_u}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{d}_j^{\tau_u}\tilde{z}_{ik_a}^{\tau_u}\right|$$

Applying the bound to each element on the right hand side of (A14), we then have

$$\text{(A22)}\quad N^{-m-\frac{\tau}{2}}\left|J(\tau,p,q,\{e_1\},...,\{e_m\})\right| \le N^{p+q-s-m}\prod_{u=1}^{s}\prod_{a=1}^{r}\overline{M}(\tau_u,v_{k_a})^{f_{au}/\tau_u}.$$

Let us now assume (to be proven later) that (A2) implies that

$$\text{(A23)}\quad N^{-\frac{\tau_u}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\widetilde{d}_j^{\tau_u}\widetilde{z}_{ik}^{\tau_u} = o(1) \quad \text{for all } k \text{ and } \tau_u = 3, 4, ... \quad ,$$

then we see that if τ_u is even and greater than 2, $\overline{M}(\tau_u, v_{k_a}) \rightarrow 0$. If τ_u is odd and greater than 1, we can apply the Cauchy-Schwarz inequality

$$\text{(A24)}\quad \overline{M}(2t+1,v_{k_a})^2 = \left(N^{-\frac{2t+1}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|v_{k_a}\right|^t\left|v_{k_a}\right|^{t+1}\right)^2$$

$$\le \left(N^{-\frac{2t}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|v_{k_a}\right|^{2t}\right)\left(N^{-\frac{2t+2}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|v_{k_a}\right|^{2t+2}\right)$$

$$= \left(N^{-\frac{2t}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}v_{k_a}^{2t}\right)\left(N^{-\frac{2t+2}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}v_{k_a}^{2t+2}\right) = o(1) \text{ for } t = 1, 2, ...$$

Finally, we have

$$\text{(A25)}\quad \overline{M}(2,n_{pq}) = N^{-2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\widetilde{d}_j^2\widetilde{z}_{ik_a}^2\right| = N^{-2}\sum_{i=1}^{N}\widetilde{d}_i^2\sum_{j=1}^{N}\widetilde{z}_{jk_a}^2 = 1.$$

Combining these results with (A22), and the fact that $p+q \le s+m$, we see that if $\tau_u \ge 2$ for all u in 1..s and (a) $\tau_u > 2$ for any u or (b) $\tau_u = 2$ for all u and $p+q < s+m$, then $N^{-m-\tau/2}J(\tau...)$ asymptotically equals 0.

We now return to the equality in (A14), expressing $J(\tau...)$ as the product of s $J(\tau_u...)$. If $\tau_u = 1$ we have $m_u = p_u = q_u = 1$, and $J(\tau_u...)$ is given by

$$\text{(A26)}\quad \sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\widetilde{d}_{j_1}\widetilde{z}_{i_1k_1} = \sum_{i_1=1}^{N}\widetilde{z}_{i_1k_1}\sum_{j_1=1}^{N}\widetilde{d}_{j_1} = 0,$$

from which it follows that $N^{-m-\tau/2}J(\tau...) = 0$ for all N. Hence, the only case where $N^{-m-\tau/2}J(\tau...)$ may not be identically or asymptotically zero is where $\tau_u = 2$ for all u. This means that each $J(\tau_u, p_u, q_u, \{e_{u1}\},...,\{e_{um_u}\})$ involves two elements, v_{k_1} and v_{k_2}, divided into $m_u = 1$ or 2 groups. If $m_u = 2$, then $p_u + q_u \le 3$. If $p_u + q_u = 3$, then $J(\tau_u...)$ is given by

$$(\text{A27})\quad \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\sum_{j_1=1}^{N}\tilde{d}_{j_1}\tilde{d}_{j_1}\tilde{z}_{i_1k_1}\tilde{z}_{i_2k_2} \;\text{ or }\; \sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\sum_{j_2=1}^{N}\tilde{d}_{j_1}\tilde{d}_{j_2}\tilde{z}_{i_1k_1}\tilde{z}_{i_1k_2}$$

both of which are zero. If $p_u + q_u = 2$ for any u, then $p + q + s - m < 0$, and by the results of the previous paragraph $N^{-m-\tau/2}J(\tau...)$ is asymptotically zero.

From the above, we see that the only case where $N^{-m-\tau/2}J(\tau...)$ may not be identically or asymptotically zero is when for each subcomponent $J(\tau_u...)$ we have $\tau_u = 2$ and $m_u = p_u = q_u = 1$ (as $p_u \leq m_u$, $q_u \leq m_u$), i.e. there is only one grouping of two v_ks (possibly the same), summed across one index for i and one for j, i.e.

$$(\text{A28})\quad J(\tau_u = 2, p_u = 1, q_u = 1, \{v_{k_1}, v_{k_2}\}) = \sum_{i_1=1}^{N}\sum_{j_1=1}^{N}\tilde{d}_{j_1}\tilde{d}_{j_1}\tilde{z}_{i_1k_1}\tilde{z}_{i_1k_2}$$

which equals N^2 if $k_1 = k_2$ and 0 otherwise. Since $J(\tau...)$ is a product of $J(\tau_u...)$, we then know that the only form of $N^{-m-\tau/2}J(\tau...)$ that is not identically or asymptotically zero is:

$$(\text{A29})\quad N^{-m-\tau/2}J(\tau, p, q, \{v_{k_1}, v_{k_1}\}, ..., \{v_{k_m}, v_{k_m}\}) \text{ with } m = p = q = \tau/2$$

$$= N^{-\tau}\sum_{i_1=1}^{N}\sum_{j_1=1}^{N}...\sum_{i_{\tau/2}=1}^{N}\sum_{j_{\tau/2}=1}^{N}\tilde{d}_{j_1}^{2}\tilde{z}_{i_1k_1}^{2}...\tilde{d}_{j_{\tau/2}}^{2}\tilde{z}_{i_{\tau/2}k_{\tau/2}}^{2} = N^{-\tau}N^{\tau} = 1.$$

As described earlier, $I(\tau, \{e_1\}, ..., \{e_m\})$ is made up of the sum and difference of $N^{-m-\tau/2}J(\tau...)$ terms, the only one of which is not identically or asymptotically zero is given in (A29). Hence, the only $I(\tau,...)$ that is not identically or asymptotically zero is that where τ is even and

$$(\text{A30})\quad I(\tau, \{e_1\}, ..., \{e_m\}) \sim N^{-m-\frac{\tau}{2}}\sum_{i_1,...,i_m}^{N}\sum_{j_1,...,j_m}^{N}\tilde{d}_{j_1}^{e_1}\tilde{z}_{i_1}^{\{e_1\}}...\tilde{d}_{j_m}^{e_m}\tilde{z}_{i_m}^{\{e_m\}}$$

$$= N^{-\frac{\tau}{2}-\frac{\tau}{2}}J(\tau, \tau/2, \tau/2, \{e_1\}, ..., \{e_{\tau/2}\}) = N^{-\tau}N^{\tau} = 1.$$

$E_{\mathbf{d}}^{\tau}$ is made up of the sum of $I(\tau,...)$ which tie the τ v_k elements (possibly repeating) into m groups through the indices i and j. To not be identically or asymptotically zero, the $I(\tau,...)$ must involve powers of 2 of each v_k, so the only asymptotically non-zero $E_{\mathbf{d}}^{\tau}$ is that where the powers to which the r unique v_k are raised, f_1, ... ,f_r , as well as $\tau = \sum f_a$, are all even. The number of ways in which f_a objects can be tied together in pairs is (f_a -1)!! (where !! denotes the double factorial). Consequently, we have shown that for all $\tau > 2$

$$\text{(A31)}\ \ E_{\mathbf{d}}^{\tau} = E_{\mathbf{d}}\left[\prod_{a=1}^{r} v_{k_a}^{f_a}\right] \rightarrow \left[\prod_{a=1}^{r}(f_a-1)!!\right] \text{ (if all } f_a \text{ even)}, \ \ =0 \ \text{(otherwise)},$$

which are the higher moments of a vector of independent mean zero standard normals.

All that remains is to show that (A2) implies (A23). Define

$$\text{(A32)}\ \ \breve{z}_{ik} = \frac{z_{ik}-m(z_{ik})}{\left(\sum_{i=1}^{N}[z_{ik}-m(z_{ik})]^2\right)^{1/2}} \ \ \& \ \ \breve{d}_i = \frac{d_i - m(d_i)}{\left(\sum_{i=1}^{N}[d_i-m(d_i)]^2\right)^{1/2}}$$

so that (A2) may be re-expressed as

$$\text{(A33)}\ \ \lim_{N\to\infty} N^{\frac{\tau}{2}-1}\sum_{i=1}^{N}\breve{z}_{ik}^{\tau}\sum_{j=1}^{N}\breve{d}_j^{\tau} = 0 \ \ \forall\ k \ \ \& \ \ \forall\ \tau = 3, 4, ...$$

If τ is even, we can equivalently say that

$$\text{(A2)}'\ \ \lim_{N\to\infty} N^{\frac{\tau}{2}-1}\sum_{i=1}^{N}\left|\breve{z}_{ik}^{\tau}\right|\left|\sum_{j=1}^{N}\left|\breve{d}_j^{\tau}\right|\right| = 0 \ \ \forall\ k.$$

However, for any odd $\tau = 2\eta+1$, we note that by Hölder's inequality

$$\text{(A34)}\ \ N^{\frac{2\eta+1}{2}-1}\sum_{i=1}^{N}\left|\breve{z}_{ik}^{2\eta+1}\right|\left|\sum_{i=1}^{N}\left|\breve{d}_i^{2\eta+1}\right|\right| \le \left(N^{\frac{2\eta+2}{2}-1}\sum_{i=1}^{N}\left|\breve{z}_{ik}^{2\eta+2}\right|\left|\sum_{i=1}^{N}\left|\breve{d}_i^{2\eta+2}\right|\right|\right)^{½}\left(N^{\frac{2\eta}{2}-1}\sum_{i=1}^{N}\left|\breve{z}_{ik}^{2\eta}\right|\left|\sum_{i=1}^{N}\left|\breve{d}_i^{2\eta}\right|\right|\right)^{½}$$

so (A2)′ in fact applies for all τ = 3, 4, ...[1] We also note that

$$\text{(A35)}\ \ \tilde{\mathbf{d}} = \mathbf{Od}\left(\frac{\mathbf{d}'\mathbf{Od}}{N}\right)^{-½} = N^{½}\breve{\mathbf{d}} \ \ \& \ \ \tilde{\mathbf{Z}} = \mathbf{OZ}\left(\frac{\mathbf{Z}'\mathbf{OZ}}{N}\right)^{-½} = N^{½}\breve{\mathbf{Z}}\mathbf{W}, \ \text{where}\ \mathbf{W} = \mathbf{D}_{\mathbf{Z'OZ}}^{½}(\mathbf{Z}'\mathbf{OZ})^{-½}.$$

The elements of **W** are asymptotically bounded as for all N sufficiently large the determinant of $\mathbf{D}_{\mathbf{Z'OZ}}^{-½}\mathbf{Z}'\mathbf{OZ}\mathbf{D}_{\mathbf{Z'OZ}}^{-½}$ is greater than some positive constant Δ, and so

$$\text{(A36)}\ \ \sum_{i=1}^{K}\sum_{j=1}^{K} w_{ij}^2 = trace(\mathbf{W}'\mathbf{W}) = trace(\mathbf{D}_{\mathbf{Z'OZ}}^{½}(\mathbf{Z}'\mathbf{OZ})^{-1}\mathbf{D}_{\mathbf{Z'OZ}}^{½}) < K\frac{K}{\Delta/K^{K-1}} = K^{K+1}/\Delta < \infty.$$

[1]When τ = 3 and η = 1, the second square root on the right-hand side of (A34) equals 1 while the first goes to 0; in all other cases both square roots on the right hand side go to zero.

To see the last, note that by the properties of the Rayleigh quotient for any positive definite matrix **A**, $\mathbf{x}'\mathbf{x}\lambda_{\min} \leq \mathbf{x}'\mathbf{A}\mathbf{x} \leq \mathbf{x}'\mathbf{x}\lambda_{\max}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues of **A**. Consequently, $\mathbf{x}'\mathbf{A}\mathbf{x}\mathbf{x}'\mathbf{A}^{-1}\mathbf{x} \leq (\mathbf{x}'\mathbf{x})^2\lambda_{\max}/\lambda_{\min}$, as the eigenvalues of $\mathbf{A}^{-1}$ are the inverse of those of **A**. Allowing **x** to equal a vector of zeros with a 1 in the i^{th} row, we see that the i^{th} diagonal element of $\mathbf{A}^{-1}$ is less than or equal to $\lambda_{\max}/\lambda_{\min}$ divided by the i^{th} diagonal element of **A**. For the K x K matrix $\mathbf{D}_{\mathbf{Z'OZ}}^{-½}\mathbf{Z}'\mathbf{OZD}_{\mathbf{Z'OZ}}^{-½}$ with determinant greater than Δ, all diagonal elements are 1, the largest eigenvalue is less than K, and the smallest eigenvalue must be greater than $\Delta/(K^{K-1})$.

With these results in mind, we complete the proof using properties of the absolute value and Hölder's inequality to show that

$$\text{(A37)} \quad \left| N^{-\frac{\tau}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{d}_j^{\tau}\breve{z}_{ik}^{\tau} \right| = \left| N^{\frac{\tau}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\breve{d}_j^{\tau}\left(\sum_{e=1}^{K}\breve{z}_{ie}w_{ek}\right)^{\tau} \right|$$

$$= \left| N^{\frac{\tau}{2}-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\breve{d}_j^{\tau}\sum_{f_1+\ldots+f_K=\tau}\frac{\tau!}{f_1!f_2!\ldots f_K!}\prod_{e=1}^{K}\breve{z}_{ie}^{f_e}w_{ek}^{f_e} \right|$$

$$\leq N^{\frac{\tau}{2}-1}\sum_{f_1+\ldots+f_K=\tau}\frac{\tau!}{f_1!f_2!\ldots f_K!}\sum_{i=1}^{N}\sum_{j=1}^{N}\prod_{e=1}^{K}\left|\breve{d}_j^{f_e}\breve{z}_{ie}^{f_e}w_{ek}^{f_e}\right|$$

$$\leq N^{\frac{\tau}{2}-1}\sum_{f_1+\ldots+f_K=\tau}\frac{\tau!}{f_1!f_2!\ldots f_K!}\prod_{e=1}^{K}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\breve{d}_j^{\tau}\breve{z}_{ie}^{\tau}w_{ek}^{\tau}\right|\right)^{f_e/\tau}$$

$$\leq \sum_{f_1+\ldots+f_K=\tau}\frac{\tau!}{f_1!f_2!\ldots f_K!}\prod_{e=1}^{K}\left(\left|w_{ek}^{\tau}\right|N^{\frac{\tau}{2}-1}\sum_{i=1}^{N}\left|\breve{z}_{ie}^{\tau}\right|\sum_{j=1}^{N}\left|\breve{d}_j^{\tau}\right|\right)^{f_e/\tau} \rightarrow 0 \text{ [by (A2)}' \text{ above]},$$

where we use the notation $\sum_{f_1+\ldots+f_K=\tau}$ to denote the summation across all sets of K non-negative integers that sum to τ.

B. Proof for the Pairs Bootstrap of (6) and Lemma in Appendix B of the paper

As we are only examining the pairs bootstrap, to simplify notation we drop the superscript p on $\boldsymbol{\delta}$. We begin by deriving obvious results to familiarize the reader with the technique used in later, more challenging, steps. Define the random variable c_{ti} as a (0,1) indicator of whether observation i is chosen in bootstrap draw t. Obviously, c_{ti} and c_{tj} are interdependent, as only one of N observations is selected on any given draw, with, for γ and ζ each equal to any positive integer, $E(c_{ti}^{\gamma}) = N^{-1}$ and $E(c_{ti}^{\gamma}c_{tj}^{\zeta}) = 0$ if $i \neq j$, but c_{ti} and c_{sj} for $s \neq t$ are independent and identically distributed for all i and j with $E(c_{ti}^{\gamma}c_{sj}^{\zeta}) = N^{-2}$. Consequently:

$$\text{(B1) } m(\delta_i) = \frac{1}{N}\sum_{i=1}^{N}\delta_i = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{N}c_{ti} = 1, \quad E(m(\delta_i)) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{N}E(c_{ti}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{N}\frac{1}{N} = 1,$$

$$E(m(\delta_i)^2) = E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{t=1}^{N}\sum_{j=1}^{N}\sum_{s=1}^{N}c_{ti}c_{sj}\right] = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{N}E(c_{ti}c_{tj}) + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s,t=1}^{N}E(c_{ti}c_{sj})$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{t=1}^{N}\frac{1}{N} + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s,t=1}^{N}\frac{1}{N^2} = \frac{N^2}{N^3} + \frac{N^3(N-1)}{N^4} = 1,$$

$$\& \;\; E(m(\delta_i)^2) - E(m(\delta_i))^2 = 0 \;\; \text{[as expected]},$$

where I use subscripted s,t to denote the summation across values of the two indices, excluding ties between them. So, $m(\delta_i) = 1$ is a constant with zero variance (proving L1p in the Lemma).

Turning to $m(\delta_i^2)$

$$\text{(B2) } E(m(\delta_i^2)) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}E(c_{t_1i}c_{t_2i}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{N}E(c_{ti}^2) + \frac{1}{N}\sum_{i=1}^{N}\sum_{t_1,t_2=1}^{N}E(c_{t_1i}c_{t_2i}) = \frac{N^2}{N^2} + \frac{N^2(N-1)}{N^3} \to 2,$$

$$E(m(\delta_i^2)^2) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}\sum_{j=1}^{N}\sum_{u_1=1}^{N}\sum_{u_2=1}^{N}E(c_{t_1i}c_{t_2i}c_{u_1j}c_{u_2j})$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}\sum_{u_1=1}^{N}\sum_{u_2=1}^{N}E(c_{t_1i}c_{t_2i}c_{u_1i}c_{u_2i}) + \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}\sum_{u_1=1}^{N}\sum_{u_2=1}^{N}E(c_{t_1i}c_{t_2i}c_{u_1j}c_{u_2j})$$

$$= \frac{1}{N}\underbrace{\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}\sum_{u_1=1}^{N}\sum_{u_2=1}^{N}E(c_{t_1i}c_{t_2i}c_{u_1i}c_{u_2i})}_{a} + \frac{(N-1)}{N}\underbrace{\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}\sum_{u_1=1}^{N}\sum_{u_2=1}^{N}E(c_{t_1i}c_{t_2i}c_{u_1j}c_{u_2j})}_{b}$$

where

$$\text{(B3) } a = \sum_{t=1}^{N}E(c_{ti}^4) + 4\sum_{t_1,t_2=1}^{N}E(c_{t_1i}c_{t_2i}^3) + 3\sum_{t_1,t_2=1}^{N}E(c_{t_1i}^2c_{t_2i}^2) + 6\sum_{t_1,t_2,t_3=1}^{N}E(c_{t_1i}^2c_{t_2i}c_{t_3i}) + \sum_{t_1,t_2,t_3,t_4=1}^{N}E(c_{t_1i}c_{t_2i}c_{t_3i}c_{t_4i})$$

$$= \frac{N}{N} + \frac{7N(N-1)}{N^2} + \frac{6N!/(N-3)!}{N^3} + \frac{N!/(N-4)!}{N^4},$$

$$b = \sum_{t_1,s_1=1}^{N}E(c_{t_1i}^2c_{s_1j}^2) + 2\sum_{t_1,t_2,s_1=1}^{N}E(c_{t_1i}c_{t_2i}c_{s_1j}^2) + \sum_{t_1,t_2,s_1,s_2=1}^{N}E(c_{t_1i}c_{t_2i}c_{s_1j}c_{s_2j}) = \frac{N(N-1)}{N^2} + \frac{2N!/(N-3)!}{N^3} + \frac{N!/(N-4)!}{N^4}.$$

As shown, the expectation denoted by "a" is calculated by considering all ways in which the four indices might, based upon the equality of their values, be tied together in one, two, three or four groups, while the expectation of "b" is similarly calculated, but with the proviso that we can ignore cases where any t index equals an s index, as the expectation then is 0 (since $i \neq j$ in "b").

Having established the technique with these simple examples, we can consider the more general expectation for any integer power $\tau \geq 2$:

$$\text{(B4) } E(m(\delta_i^\tau)) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i}) = \sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i})$$

$$= \underbrace{\sum_{t=1}^{N} E(c_{ti}^\tau) + \sum_{a_1+a_2=\tau}\sum_{t_1,t_2=1}^{N} E(c_{t_1 i}^{a_1} c_{t_2 i}^{a_2}) + \sum_{a_1+a_2+a_3=\tau}\sum_{t_1,t_2,t_3=1}^{N} E(c_{t_1 i}^{a_1} c_{t_2 i}^{a_2} c_{t_3 i}^{a_3})... \sum_{t_1,t_2,...,t_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i})}_{c_N(\tau)}$$

$$\& \;\; c_N(\tau) = \frac{N}{N} + \sum_{a_1+a_2=\tau}\frac{N(N-1)}{N^2} + \sum_{a_1+a_2+a_3=\tau}\frac{N!/(N-3)!}{N^3} + ... + \frac{N!/(N-\tau)!}{N^\tau} = \sum_{j=1}^{\tau}\frac{N!/(N-j)!}{N^j} C_j^\tau,$$

where we use the notation $\sum_{a_1+...+a_j=\tau}$ to denote the summation across all ways in which τ objects can be divided into j groups and C_j^τ to denote the number of ways this can be achieved (as all such objects have the same expectation), with $C_1^\tau = C_\tau^\tau = 1$. Similarly, we have

$$\text{(B5) } E(m(\delta_i^\tau)^2) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N}\sum_{j=1}^{N}\sum_{s_1=1}^{N}...\sum_{s_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i} c_{s_1 j}...c_{s_\tau j})$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N}\sum_{s_1=1}^{N}...\sum_{s_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i} c_{s_1 i}...c_{s_\tau i}) + \frac{1}{N^2}\sum_{i,j=1}^{N}\sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N}\sum_{s_1=1}^{N}...\sum_{s_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i} c_{s_1 j}...c_{s_\tau j})$$

$$= \frac{1}{N}\underbrace{\sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N}\sum_{s_1=1}^{N}...\sum_{s_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i} c_{s_1 i}...c_{s_\tau i})}_{d_N(\tau)} + \frac{N-1}{N}\underbrace{\sum_{t_1=1}^{N}...\sum_{t_\tau=1}^{N}\sum_{s_1=1}^{N}...\sum_{s_\tau=1}^{N} E(c_{t_1 i}...c_{t_\tau i} c_{s_1 j}...c_{s_\tau j})}_{e_N(\tau)}$$

where, keeping in mind that in "e_N" the expectation of any object with a tie between an s and t index is zero,

$$\text{(B6) } d_N(\tau) = \sum_{j=1}^{2\tau} \frac{N!/(N-j)!}{N^j} C_j^{2\tau},$$

$$e_N(\tau) = \sum_{s,t=1}^{N} E(c_{ti}^{\tau})E(c_{sj}^{\tau}) + \sum_{a_1+a_2=\tau} \sum_{t_1,t_2,s=1}^{N} E(c_{t_1 i}^{a_1} c_{t_2 i}^{a_2})E(c_{sj}^{\tau}) + \ldots \sum_{t_1,t_2,\ldots,t_\tau,s=1}^{N} E(c_{t_1 i} \ldots c_{t_\tau i})E(c_{sj}^{\tau})$$

$$+ \sum_{b_1+b_2=\tau} \sum_{t,s_1,s_2=1}^{N} E(c_{ti}^{\tau})E(c_{t_1 i}^{b_1} c_{t_2 i}^{b_2}) + \sum_{a_1+a_2=\tau} \sum_{b_1+b_2=\tau} \sum_{t_1,t_2,s_1,s_2=1}^{N} E(c_{t_1 i}^{a_1} c_{t_2 i}^{a_2})E(c_{s_1 i}^{b_1} c_{s_2 i}^{b_2}) + \ldots \sum_{b_1+b_2=\tau} \sum_{t_1,t_2,\ldots,t_\tau,s_1,s_2=1}^{N} E(c_{t_1 i} \ldots c_{t_\tau i})E(c_{s_1 i}^{b_1} c_{s_2 i}^{b_2})$$

$$+ \ldots \quad = \sum_{j=1}^{\tau} \sum_{k=1}^{\tau} \frac{N!/(N-j-k)!}{N^{j+k}} C_j^{\tau} C_k^{\tau}.$$

We note that

$$\text{(B7) } e_N(\tau) - c_N(\tau)^2 = \sum_{j=1}^{\tau} \sum_{k=1}^{\tau} \frac{N!}{N^{j+k}(N-j-k)!} C_j^{\tau} C_k^{\tau} - \sum_{j=1}^{\tau} \sum_{k=1}^{\tau} \frac{N!N!}{N^{j+k}(N-j)!(N-k)!} C_j^{\tau} C_k^{\tau}$$

$$= \sum_{j=1}^{\tau} \sum_{k=1}^{\tau} \frac{N! C_j^{\tau} C_k^{\tau}}{N^{j+k}(N-j-k)!} \left[1 - \frac{N(N-1)\ldots(N-j+1)}{(N-k)(N-k+1)\ldots(N-k-j+1)} \right] < 0.$$

Using the above, we see that $m(\delta_i^{\tau})$ converges in mean square and hence in probability to the value given in (B4) as its variance is seen to be $O(N^{-1})$

$$\text{(B8) } E[m(\delta_i^{\tau}) - E[m(\delta_i^{\tau})]]^2 = E[m(\delta_i^{\tau})^2] - E[m(\delta_i^{\tau})]^2$$

$$= \frac{d_N(\tau)}{N} + \frac{(N-1)}{N} e_N(\tau) - c_N(\tau)^2 < \frac{d_N(\tau)}{N} < \frac{\sum_{j=1}^{2\tau} C_j^{2\tau}}{N}.$$

As (B4) converges to a finite constant

$$\text{(B9) } \sum_{j=1}^{\tau} \frac{N!/(N-j)!}{N^j} C_j^{\tau} \to \sum_{j=1}^{\tau} C_j^{\tau},$$

this proves (L3b). As this applies to $\tau = 2$ as well, using (B1) and (B2) we prove (L2b) and (6) in the paper:

$$\text{(B10) } \frac{\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}}{N} = m(\delta_i^2) - m(\delta_i) \overset{p(\boldsymbol{\delta})}{\to} 2 - 1 = 1.$$

C. Consistency of the Wild Bootstrap without Higher Moments on δ_i^w

This appendix proves consistency of the wild bootstrap without higher moments on δ_i^w. Specifically, we only assume that $E[\delta_i^w] = 0$, $E[(\delta_i^w)^2] = 1$ & $E[(\delta_i^w)^{2(1+\theta_1)}] < \Delta$ for some finite Δ

and $\theta_1 \geq \max(1, 1/\gamma_2+\kappa, 1/\gamma_3+\kappa)$, for κ some constant > 0 and γ_2 & γ_3 as given in Theorem V. All of White's (1980) assumptions in Theorem V are assumed to hold. As we are concentrating only on the wild bootstrap, we also strengthen the proof to almost surely across $\mathbf{X},\boldsymbol{\varepsilon}$. To simplify notation we drop the superscript w on δ_i^w. We make use of the Markov corollary and Borel-Cantelli corollary given in Appendix C of the paper. We also note that White's (1984) Corollary to the Continuous Mapping Theorem also holds for convergence almost surely.

We begin by showing that White's (1980) assumptions actually imply stronger conditions than (IVa) - (IVd), namely that any θ in $(0, \min(\gamma_2/(1+\gamma_2), \gamma_3/(1+\gamma_3)))$:

$$\text{(C1a)} \quad \frac{\mathbf{X}'\mathbf{X}}{N} - \mathbf{V}_{N_1} \overset{a.s.(\mathbf{X})}{\to} \mathbf{0}_{KxK} \quad \& \quad \frac{\mathbf{X}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{X}}{N} - \mathbf{V}_{N_2} \overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} \mathbf{0}_{KxK}$$

$$\text{where } \mathbf{V}_{N_i} \ (i=1,2) \text{ is bounded and non-singular with}$$
$$\text{determinant}(\mathbf{V}_{N_i}) > \eta_i > 0 \text{ for all } N \text{ sufficiently large};;$$

$$\text{(C1b)} \quad \forall\ k \,\&\, \tau: \frac{N^{\theta\left(\frac{\tau}{2}-1\right)}\sum_{i=1}^{N} x_{ik}^{\tau}\hat{\varepsilon}_i^{\tau}}{\left(\sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2}} \overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0;$$

$$\text{(C1c)} \ \forall\ k: \sum_{i=1}^{N}\frac{x_{ik}^4}{N^{2-\theta}} \overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0; \quad \text{(C1d)} \ \forall\ k: \sum_{i=1}^{N}\frac{x_{ik}^8}{N^{4-3\theta}} \overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0.$$

Appendix C in the paper proved (C1a). We note that from the Corollary to the Continuous Mapping Theorem this implies that almost surely $\mathbf{X}'\mathbf{X}/N$ & $\mathbf{X}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{X}/N$ are invertible with determinants strictly greater than some $\eta > 0$.

Continuing, we bound (C1b) - (C1d)

$$\text{(C2a)} \quad \left|N^{\theta\left(\frac{\tau}{2}-1\right)}\sum_{i=1}^{N} x_{ik}^{\tau}\hat{\varepsilon}_i^{\tau}\right| \Big/ \left(\sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2} \leq \left(\operatorname*{Max}_i x_{ik}^2\hat{\varepsilon}_i^2 / N^{1-\theta} \Big/ \sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2 / N\right)^{\frac{\tau}{2}-1}$$

$$\text{(C2b)} \quad \sum_{i=1}^{N}\frac{x_{ik}^4}{N^{2-\theta}} \leq \frac{\operatorname*{Max}_i x_{ik}^2}{N^{1-\theta}}\sum_{i=1}^{N}\frac{x_{ik}^2}{N} \quad \& \quad \sum_{i=1}^{N}\frac{x_{ik}^8}{N^{4-3\theta}} \leq \left(\frac{\operatorname*{Max}_i x_{ik}^2}{N^{1-\theta}}\right)^3\sum_{i=1}^{N}\frac{x_{ik}^2}{N},$$

so, to prove (C1b) - (C1d) it is sufficient to show that the right hand sides of the inequalities above converge to zero. From the Markov corollary and (Va) and (Vb) in Theorem V, we know that $\sum_{i=1}^{N} x_{ik}^2 / N$ converges almost surely to $\sum_{i=1}^{N} E(x_{ik}^2)/N$, which is bounded from above. Using Markov's inequality and the assumption that $E(|x_{ij}^2|^{1+\gamma_1}) < \Delta_1$ in (Vb) of Theorem V, we can state that for some $\gamma > 1/(1+\gamma_1)$ but < 1

$$\text{(C3)} \quad \sum_{i=1}^{\infty} P(x_{ik}^2 \geq N^{\gamma}) \leq \sum_{i=1}^{\infty} \frac{E(|x_{ik}^2|^{1+\gamma_1})}{N^{\gamma(1+\gamma_1)}} < \sum_{i=1}^{\infty} \frac{\Delta_1}{N^{\gamma(1+\gamma_1)}} < \infty.$$

By the Borel-Cantelli corollary, $\text{Max}_i x_{ik}^2$ is asymptotically almost surely less than N^{γ} and hence $\text{Max}_i x_{ik}^2 / N^{1-\theta}$ almost surely converges to zero for θ in (0, $\gamma_1/(1+\gamma_1)$) [implied by $1-\theta > \gamma > (1+\gamma_1)^{-1}$], which in conjunction with the almost sure convergence of $\sum_{i=1}^{N} x_{ik}^2 / N$ to a bounded value establishes that the expressions in (C2b) converge to zero and hence (C1c) and (C1d) hold.

Turning to (C2a), as noted in Appendix C of the paper White's assumptions ensure that the term $\sum_{i=1}^{N} x_{ik}^2 \hat{\varepsilon}_i^2 / N$ in the denominator of (C2a) is almost surely bounded from below away from zero. Regarding the max term, from (C3) in Appendix C of the paper we see that to prove that $\text{Max}_i x_{ik}^2 \hat{\varepsilon}_i^2 / N^{1-\theta}$ converges almost surely to zero it is sufficient to show that $\text{Max}_i x_{ik}^2 \varepsilon_i^2 / N^{1-\theta}$ and $\text{Max}_i x_{ik}^4 / N^{1-\theta}$ converge almost surely to zero for all *k*. However, the assumptions $E(|\varepsilon_i^2 x_{ij} x_{ik}|^{1+\gamma_2}) < \Delta_2$ and $E(|x_{ij}^4|^{1+\gamma_3}) < \Delta_3$ in (Vc) and (Vd) of Theorem V, by the same argument used in (C3) above, ensure that this is the case for θ in (0,min($\gamma_2/(1+\gamma_2)$, $\gamma_3/(1+\gamma_3)$)), establishing (IVb). Earlier above, θ was seen to lie in (0,$\gamma_1/(1+\gamma_1)$), but as, by their definitions in Theorem V, $\gamma_1 = 1+2\gamma_3$, these bounds on θ are tighter. In sum, White's assumptions ensure that (C1a) - (C1d) hold for all θ greater than 0 and less than the minimum of $\gamma_2/(1+\gamma_2)$ and $\gamma_3/(1+\gamma_3)$.

Continuing, from Theorems I and II in the paper the distribution across **d** of

$$\text{(C4)} \quad \left(\frac{\mathbf{X}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{X}}{N}\right)^{-1/2} \left(\frac{\mathbf{d}'\mathbf{O}\mathbf{d}}{N}\right)^{-1/2} \frac{\mathbf{X}'\mathbf{D}_{\mathbf{d}}\hat{\boldsymbol{\varepsilon}}}{\sqrt{N}}$$

converges almost surely (across $\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon}$) to that of the iid multivariate standard normal provided for all N sufficiently large $\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}'$ is non-zero and the correlation matrix $\mathbf{D}_{\mathbf{Z}'\mathbf{OZ}}^{-1/2}\mathbf{Z}'\mathbf{OZD}_{\mathbf{Z}'\mathbf{OZ}}^{-1/2}$ is non-singular with determinant > Δ > 0, and for all integer τ greater than 2 and some θ in (0,1)

$$\text{(C5)}\quad \left| \frac{N^{\tau/2-1}\sum_{i=1}^{N} x_{ik}^{\tau}\hat{\varepsilon}_i^{\tau}}{\left(\sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2}} \frac{\sum_{i=1}^{N}[\delta_i - m(\delta_i)]^{\tau}}{\left(\sum_{i=1}^{N}[\delta_i - m(\delta_i)]^2\right)^{\tau/2}} \right| \leq \frac{N^{\tau/2-1}\sum_{i=1}^{N}\left|x_{ik}^{\tau}\hat{\varepsilon}_i^{\tau}\right|}{\left(\sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2}} \frac{\sum_{i=1}^{N}\left|[\delta_i - m(\delta_i)]^{\tau}\right|}{\left(\sum_{i=1}^{N}[\delta_i - m(\delta_i)]^2\right)^{\tau/2}}$$

$$\leq \frac{N^{\tau/2-1}(\operatorname*{Max}_i x_{ik}^2\hat{\varepsilon}_i^2)^{\tau/2-1}(\operatorname*{Max}_i[\delta_i - m(\delta_i)]^2)^{\tau/2-1}\sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2\sum_{i=1}^{N}[\delta_i - m(\delta_i)]^2}{\left(\sum_{i=1}^{N} x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2}\left(\sum_{i=1}^{N}[\delta_i - m(\delta_i)]^2\right)^{\tau/2}}$$

$$= \frac{\left(\operatorname*{Max}_i \dfrac{x_{ik}^2\hat{\varepsilon}_i^2}{N^{1-\theta}}\right)^{\tau/2-1}\left(\operatorname*{Max}_i \dfrac{[\delta_i - m(\delta_i)]^2}{N^{\theta}}\right)^{\tau/2-1}}{\left(\sum_{i=1}^{N}\dfrac{x_{ik}^2\hat{\varepsilon}_i^2}{N}\right)^{\tau/2-1}\left(\sum_{i=1}^{N}\dfrac{[\delta_i - m(\delta_i)]^2}{N}\right)^{\tau/2-1}} \overset{a.s.(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow} 0.$$

From the assumptions $E[\delta_i^w]=0$ and $E[(\delta_i^w)^2]=1$ and the Strong Law of Large Numbers, we know that $\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}'/N$ almost surely converges to 1, and hence for all N sufficiently large $\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}'$ is almost surely non-zero, while White's Corollary to the Continuous Mapping Theorem and the argument in footnote 7 in the paper ensure that for all N sufficiently large $\mathbf{D}_{\mathbf{Z}'\mathbf{OZ}}^{-1/2}\mathbf{Z}'\mathbf{OZD}_{\mathbf{Z}'\mathbf{OZ}}^{-1/2}$ is almost surely non-singular with determinant > Δ > 0. Turning to (C5), we have already seen that $\text{Max}_i x_{ik}^2\hat{\varepsilon}_i^2/N^{1-\theta}$ converges almost surely to 0 for any θ in (0,min(γ_2/(1+ γ_2), γ_3/(1+ γ_3))). As, following Markov's Inequality, for θ > 1/(1+θ_1) (such that $E[\delta_i^{2(1+\theta_1)}] < \Delta$)

$$\text{(C6)}\quad \sum_{i=1}^{\infty} P(\delta_i^2 \geq N^{\theta}) \leq \sum_{i=1}^{\infty}\frac{E(|\delta_i^2|^{(1+\theta_1)})}{N^{\theta(1+\theta_1)}} < \sum_{i=1}^{\infty}\frac{\Delta}{N^{\theta(1+\theta_1)}} < \infty$$

$$\&\quad \sum_{i=1}^{\infty} P(|\delta_i| \geq N^{\theta}) \leq \sum_{i=1}^{\infty}\frac{E(|\delta_i|^{2(1+\theta_1)})}{N^{2\theta(1+\theta_1)}} < \sum_{i=1}^{\infty}\frac{\Delta}{N^{2\theta(1+\theta_1)}} < \infty,$$

by a similar application of the Borel-Cantelli Lemma given in Appendix C and using the fact that by the Strong Law of Large Numbers $m(\delta_i) \overset{a.s.(\boldsymbol{\delta})}{\rightarrow} 0$, we see that

$$\text{(C7)}\ \underset{i}{\text{Max}}\frac{[\delta_i - m(\delta_i)]^2}{N^{\theta}} \leq \underset{i}{\text{Max}}\frac{\delta_i^2}{N^{\theta}} + m(\delta_i)\underset{i}{\text{Max}}\frac{2|\delta_i|}{N^{\theta}} + \frac{m(\delta_i)^2}{N^{\theta}}$$

is almost surely bounded. Consequently, (C5) holds if $\theta_1 > \max(1/\gamma_2, 1/\gamma_3)$ [using $\theta_1 > 1/\theta - 1$]. Following (B6) in Appendix B of the paper, we then see that $\sqrt{N}(\hat{\boldsymbol{\beta}}^w_{\mathbf{d},N} - \hat{\boldsymbol{\beta}}_N)$ converges to the multivariate mean zero normal with covariance matrix $N\mathbf{V}(\hat{\boldsymbol{\beta}}_N)$.

Next we turn to the wild bootstrap covariance estimates given in (B8) in the paper. Results (B10), (B11), (B13) and (B14) in the paper are seen, given (C1) above, to hold almost surely. Consequently, the almost sure fulfillment of Condition (IIIa) for "*a*" and "*b*" in (B8) requires that $m(\delta_i^4) - m(\delta_i^2)^2$ & $m(\delta_i^2) - m(\delta_i)^2$ are almost surely bounded. For this, it suffices that $E[(\delta_i^w)^4]$ is finite. This then ensures that $N\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}^w_{\mathbf{d},N}) - N\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_N) \xrightarrow{p(\mathbf{d})|a.s.(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})} \mathbf{0}_{KxK}$ as in (B17) in the paper. Together, these results show $E[(\delta_i^w)^{2(1+\theta_1)}] < \Delta < \infty$, for $\theta_1 \geq \max(1, 1/\gamma_2+\kappa, 1/\gamma_3+\kappa)$ and some $\kappa > 0$, is sufficient to ensure the almost sure consistency of the wild bootstrap.

D. Consistency of the Pairs Bootstrap with Sub-Sampling

This appendix proves consistency of the pairs bootstrap with sub-sampling $M < N$ observations, with and without replacement. As noted in the paper, we assume that for some $\gamma^* > \max((1+\gamma_2)^{-1}, (1+\gamma_3)^{-1}))$ $\exists c > 0$ and $\exists N_0$ such that $\forall N > N_0$ $M > cN^{\gamma^*}$, i.e. $\lim_{\inf} M/N^{\gamma^*} > 0$. We also assume that $\lim_{\sup} M/N < 1$ and in particular for sampling without replacement that $M/N \to 0$. We modify the notation, so that $\boldsymbol{\Delta}$ is now an M x N matrix of 0s with a single 1 in each row. Otherwise, the notation is as before, with the bootstrap sample given by $\boldsymbol{\Delta}(\mathbf{y},\mathbf{X})$ and the associated estimated coefficients, residuals and covariance matrix:

$$\text{(D1)}\ \hat{\boldsymbol{\beta}}^p_{\boldsymbol{\delta}^p,N} = (\mathbf{X}'\mathbf{D}_{\boldsymbol{\delta}^p}\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}_{\boldsymbol{\delta}^p}\mathbf{y} = \hat{\boldsymbol{\beta}}_N + (\mathbf{X}'\mathbf{D}_{\boldsymbol{\delta}^p}\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}_{\boldsymbol{\delta}^p}\hat{\boldsymbol{\varepsilon}}_N$$

$$\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{\delta}^p,N} = \boldsymbol{\Delta}\hat{\boldsymbol{\varepsilon}}_N + \boldsymbol{\Delta}\mathbf{X}(\hat{\boldsymbol{\beta}}_N - \hat{\boldsymbol{\beta}}^p_{\boldsymbol{\delta}^p,N}),\ \text{and}\ \hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}^p,N}) = (\mathbf{X}'\mathbf{D}_{\boldsymbol{\delta}^p}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Delta}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{\delta}^p,N}}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{\delta}^p,N}}\boldsymbol{\Delta}\mathbf{X}(\mathbf{X}'\mathbf{D}_{\boldsymbol{\delta}^p}\mathbf{X})^{-1},$$

As we are only considering the pairs bootstrap in this appendix, below we drop the superscript *p* on $\boldsymbol{\delta}$. We wish to show that:

(D2a) $\left(\frac{\mathbf{X}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{X}}{N}\right)^{-½}\left(\frac{\mathbf{X}'\mathbf{X}}{N}\right)\left(\frac{\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}}{M}\right)^{-½}\sqrt{M}(\hat{\boldsymbol{\beta}}^{b}_{\mathbf{d},N}-\hat{\boldsymbol{\beta}}_N)\overset{d(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}\mathbf{n}_K$

(D2b) $M\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}^{b}_{\mathbf{d},N})\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}N\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_N),$

which, along with the result that $\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}/M\overset{p(\boldsymbol{\delta})}{\rightarrow}1$ in the Lemma below, following the arguments given in section IV of the paper is enough to ensure conditional consistency of coefficient estimates and Wald statistics.

We know that White's (1980) assumptions are sufficient to ensure that (C1a) - (C1d) in Appendix C above hold for all θ in $(0,\min(\gamma_2/(1+\gamma_2), \gamma_3/(1+\gamma_3)))$. We apply (C1a) - (C1d) below assuming that θ lies in $(1-\gamma^*,\min(\gamma_2/(1+\gamma_2), \gamma_3/(1+\gamma_3)))$, so that $1 - \theta - \gamma^* < 0$. In this case

$$\text{(D3)}\quad \frac{N}{MN^{\theta}}=\frac{N^{1-\theta-\gamma^*}}{M/N^{\gamma^*}}=O(N^{1-\theta-\gamma^*})\rightarrow 0.$$

The following Lemma, proven at the end of this appendix, will be useful:

Lemma: Let $\overset{p(\boldsymbol{\delta})}{\rightarrow}$ denote convergence in probability across the distribution of $\boldsymbol{\delta}$, τ any integer greater than or equal to 2, and θ a constant in $(1-\gamma^*,\min(\gamma_2/(1+\gamma_2), \gamma_3/(1+\gamma_3)))$. Then:

$$\text{(L1)}\ m(\delta_i)=\frac{M}{N};\quad \text{(L2)}\left(\frac{N}{MN^{\theta}}\right)\frac{N}{M}m(\delta_i^2)\overset{p(\boldsymbol{\delta})}{\rightarrow}0;$$

$$\text{(L3)}\ \frac{\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^{\tau}}{N^{1+\theta\left(\frac{\tau}{2}-1\right)}}}{\left(\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^2}{N}\right)^{\tau/2}}\overset{p(\boldsymbol{\delta})}{\rightarrow}0;\ \text{(L4)}\sum_{i=1}^{N}[\delta_i-m(\delta_i)]^2>0;\ \text{(L5)}\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^2}{M}\overset{p(\boldsymbol{\delta})}{\rightarrow}1.$$

Starting with the pairs bootstrap coefficient estimates, we have:

$$\text{(D4)}\ \sqrt{M}(\hat{\boldsymbol{\beta}}^{p}_{\mathbf{d},N}-\hat{\boldsymbol{\beta}}_N)=\mathbf{C}^{-1}\mathbf{a},\ \text{where}\ \mathbf{C}=\frac{\mathbf{X}'\mathbf{D}_{\mathbf{d}}\mathbf{X}(N/M)}{N}\ \text{and}\ \mathbf{a}=\frac{\sqrt{(N/M)}\mathbf{X}'\mathbf{D}_{\mathbf{d}}\hat{\boldsymbol{\varepsilon}}}{\sqrt{N}}.$$

Regarding the jk^{th} element of $\mathbf{C}$, given by $\sum_{i=1}^{N}x_{ij}x_{ik}(N/M)d_i/N$, we can apply Theorem III with $z_i=x_{ij}x_{ik}$ and "d_i" = $(N/M)d_i$. As $m(x_{ij}x_{ik})$ is almost surely bounded by (C1a) and

$$\text{(D5)}\ \sum_{i=1}^{N}\frac{x_{ij}^2 x_{ik}^2}{N^{2-\theta}} \le \sqrt{\sum_{i=1}^{N}\frac{x_{ij}^4}{N^{2-\theta}}\sum_{i=1}^{N}\frac{x_{ik}^4}{N^{2-\theta}}}\ \underbrace{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}}_{\text{(C1c)}}\ 0$$

we see that Condition IIIa is satisfied

$$\text{(D6)}\ \underbrace{\frac{m(x_{ij}^2x_{ik}^2)-m(x_{ij}x_{ik})^2}{N^{1-\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}\ 0\ \text{[C1a and C1c]}}\left(\underbrace{\frac{N}{MN^{\theta}}\frac{N}{M}m(\delta_i^2)}_{\overset{p(\boldsymbol{\delta})}{\rightarrow}\ 0\,\text{(L2)}}-\underbrace{\frac{N^2}{M^2N^{\theta}}m(\delta_i)^2}_{\rightarrow 0\,\text{(L1)}}\right)\overset{p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0,$$

and so

$$\text{(D7)}\ \underbrace{\frac{\mathbf{X}'\mathbf{D}_{\mathbf{d}}\mathbf{X}(N/M)}{N}}_{\mathbf{C}}-\frac{\mathbf{X}'\mathbf{X}}{N}\underbrace{\frac{N}{M}m(\delta_i)}_{=1\,\text{(L1)}}\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}\mathbf{0}_{KxK}.$$

Noting that the k^{th} element of **a** equals $\sum_{i=1}^{N}x_{ik}\hat{\varepsilon}_i d_i\sqrt{N/M}/\sqrt{N}$, we apply the multivariate extension of Theorem I in the text with $z_{ik}=x_{ik}\hat{\varepsilon}_i$, or $\mathbf{Z}=\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{X}$ and the "d_i" = $d_i\sqrt{N/M}$. From (L4) we know that $\mathbf{d}'\mathbf{O}\mathbf{d}=\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}$ is non-zero, while assumption (C1a) ensures that $\mathbf{D}^{-1/2}_{\mathbf{X}'\mathbf{D}_{\hat{\varepsilon}}\mathbf{D}_{\hat{\varepsilon}}\mathbf{X}}\mathbf{X}'\mathbf{D}_{\hat{\varepsilon}}\mathbf{D}_{\hat{\varepsilon}}\mathbf{X}\mathbf{D}^{-1/2}_{\mathbf{X}'\mathbf{D}_{\hat{\varepsilon}}\mathbf{D}_{\hat{\varepsilon}}\mathbf{X}}$ is almost surely non-singular with determinant greater than some $\Delta > 0$ [see Appendix C above]. Hence, following Theorems I and II, the distribution across **d** of

$$\text{(D8)}\left(\frac{\mathbf{X}'\mathbf{D}_{\hat{\varepsilon}}\mathbf{D}_{\hat{\varepsilon}}\mathbf{X}}{N}\right)^{-1/2}\left(\frac{(N/M)\mathbf{d}'\mathbf{O}\mathbf{d}}{N}\right)^{-1/2}\frac{\sqrt{(N/M)}\mathbf{X}'\mathbf{D}_{\mathbf{d}}\hat{\boldsymbol{\varepsilon}}}{\sqrt{N}}$$

converges in probability (across **δ**,**X**,**ε**) to that of the iid multivariate standard normal provided that for all integer τ greater than 2

$$\text{(D9)}\ \frac{N^{\theta\left(\frac{\tau}{2}-1\right)}\sum_{i=1}^{N}x_{ik}^{\tau}\hat{\varepsilon}_i^{\tau}}{\left(\sum_{i=1}^{N}x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2}}\frac{\sum_{i=1}^{N}\frac{\left[\sqrt{\frac{N}{M}}\delta_i-m(\sqrt{\frac{N}{M}}\delta_i)\right]^{\tau}}{N^{1+\theta\left(\frac{\tau}{2}-1\right)}}}{\left(\sum_{i=1}^{N}\frac{\left[\sqrt{\frac{N}{M}}\delta_i-m(\sqrt{\frac{N}{M}}\delta_i)\right]^2}{N}\right)^{\tau/2}}=\frac{N^{\theta\left(\frac{\tau}{2}-1\right)}\sum_{i=1}^{N}x_{ik}^{\tau}\hat{\varepsilon}_i^{\tau}}{\left(\sum_{i=1}^{N}x_{ik}^2\hat{\varepsilon}_i^2\right)^{\tau/2}}\frac{\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^{\tau}}{N^{1+\theta\left(\frac{\tau}{2}-1\right)}}}{\left(\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^2}{N}\right)^{\tau/2}}\overset{p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0,$$

which given (L3) as well as (C1b) in Appendix C is satisfied. Putting the preceding together, we have:

$$\text{(D10)}\ \left(\frac{\mathbf{X'D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{X}}{N}\right)^{-½}\left(\frac{\mathbf{X'X}}{N}\right)\left(\frac{\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}}{M}\right)^{-½}\sqrt{M}\,(\hat{\boldsymbol{\beta}}^{b}_{\mathbf{d},N}-\hat{\boldsymbol{\beta}}_N)=$$

$$\left(\frac{\mathbf{X'D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_N}\mathbf{X}}{N}\right)^{-½}\left(\frac{\mathbf{X'X}}{N}\right)\underbrace{\left(\frac{(N/M)\mathbf{X'D_dX}}{N}\right)^{-1}}_{\substack{p|(\mathbf{d})p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})\\ \to \mathbf{X'X}/N}}\left(\frac{\mathbf{X'D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{X}}{N}\right)^{½}\underbrace{\left(\frac{\mathbf{X'D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{X}}{N}\right)^{-½}\left(\frac{(N/M)\mathbf{d'Od}}{N}\right)^{-½}\frac{\sqrt{\frac{N}{M}}\mathbf{X'D_d}\hat{\boldsymbol{\varepsilon}}}{\sqrt{N}}}_{\substack{d(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})\\ \to\ \mathbf{n}_K}},$$

$$\overset{d(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\to}\ \mathbf{n}_K$$

thereby establishing the claim in (D2a).

Regarding the pairs bootstrap heteroskedasticity robust covariance estimates, we have

$$\text{(D11)}\ M\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}^{p}_{\mathbf{d},N})=\left(\frac{\mathbf{X'D_dX}(N/M)}{N}\right)^{-1}\mathbf{B}\left(\frac{\mathbf{X'D_dX}(N/M)}{N}\right)^{-1},\ \text{where}\ \mathbf{B}=\frac{(N/M)\mathbf{X}'\boldsymbol{\Delta}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_{\mathbf{d},N}}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}_{\mathbf{d},N}}\boldsymbol{\Delta}\mathbf{X}}{N}.$$

The jk^{th} element of **B** is given by

$$\text{(D12)}\ \sum_{i=1}^{N}\frac{\frac{N}{M}x_{ij}x_{ik}d_i(\hat{\varepsilon}_i-\sum_{p=1}^{K}\frac{x_{ip}}{\sqrt{M}}\gamma_N^{½}\hat{\eta}_p)^2}{N},\left[\text{where}\ \hat{\boldsymbol{\eta}}=\gamma_N^{-½}\sqrt{M}\,(\hat{\boldsymbol{\beta}}^{p}_{\mathbf{d},N}-\hat{\boldsymbol{\beta}}_N)\ \&\ \gamma_N=\frac{\boldsymbol{\delta}'\mathbf{O}\boldsymbol{\delta}}{M}\right]$$

$$=\underbrace{m\left(\frac{N}{M}x_{ij}x_{ik}\hat{\varepsilon}_i^2d_i\right)}_{d}-2\sum_{p=1}^{K}\gamma_N^{½}\hat{\eta}_p\underbrace{m\left(\frac{N}{M^{3/2}}x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_id_i\right)}_{e}+\sum_{p=1}^{K}\sum_{q=1}^{K}\gamma_N\hat{\eta}_p\hat{\eta}_q\underbrace{m\left(\frac{N}{M^2}x_{ij}x_{ik}x_{ip}x_{iq}d_i\right)}_{f}.$$

For "d", we apply Theorem III with $z_i=x_{ij}x_{ik}\hat{\varepsilon}_i^2$ and "d_i" = $d_i(N/M)$. As $m(x_{ij}x_{ik}\hat{\varepsilon}_i^2)$ is almost surely bounded by (C1a) and

$$\text{(D13)}\ \frac{m(x_{ij}^2x_{ik}^2\hat{\varepsilon}_i^4)}{N^{1-\theta}}=\sum_{i=1}^{N}\frac{x_{ij}^2x_{ik}^2\hat{\varepsilon}_i^4}{N^{2-\theta}}\le\sqrt{\left(\sum_{i=1}^{N}\frac{x_{ij}^4\hat{\varepsilon}_i^4}{N^{2-\theta}}\right)\left(\sum_{i=1}^{N}\frac{x_{ik}^4\hat{\varepsilon}_i^4}{N^{2-\theta}}\right)}$$

$$\text{with}\ \sum_{i=1}^{N}\frac{x_{ij}^4\hat{\varepsilon}_i^4}{N^{2-\theta}}=\frac{\overbrace{N^{\theta}\sum_{i=1}^{N}(x_{ij}\hat{\varepsilon}_i)^4}^{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to}0\ \text{(C1b with }\tau=4)}}{\left(\sum_{i=1}^{N}x_{ij}^2\hat{\varepsilon}_i^2\right)^{4/2}}\overbrace{\left(\sum_{i=1}^{N}\frac{x_{ij}^2\hat{\varepsilon}_i^2}{N}\right)^2}^{\text{almost surely bounded (C1a)}}\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to}0.$$

Condition IIIa is met

$$\text{(D14)}\ \underbrace{\frac{m(x_{ij}^2x_{ik}^2\hat{\varepsilon}_i^4)-m(x_{ij}x_{ik}\hat{\varepsilon}_i^2)^2}{N^{1-\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}\ 0\ (\text{C1a \& D13})}\left(\underbrace{\frac{N}{MN^{\theta}}\frac{N}{M}m(\delta_i^2)}_{\overset{p(\boldsymbol{\delta})}{\rightarrow}0\,(\text{L2})}-\underbrace{\frac{N^2}{M^2N^{\theta}}m(\delta_i)^2}_{\rightarrow 0\,(\text{L1})}\right)\overset{p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0,$$

and so

$$\text{(D15) "}d\text{": } m(x_{ij}x_{ik}\hat{\varepsilon}_i^2 d_i(N/M))-m(x_{ij}x_{ik}\hat{\varepsilon}_i^2)\underbrace{\frac{N}{M}m(\delta_i)}_{=1\,(\text{L1})}\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0.$$

For "*e*", we apply Theorem III with $z_i = x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_i/\sqrt{M}$ and "d_i" = $d_i(N/M)$. In this case we note that

$$\text{(D16)}\left|m\left(\frac{x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_i}{\sqrt{M}}\right)\right|\le\sum_{i=1}^{N}\frac{\left|x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_i\right|}{N^{(3-\theta)/2}}\sqrt{\frac{N}{MN^{\theta}}}\le\underbrace{\sqrt{\frac{N}{MN^{\theta}}}}_{\rightarrow 0}\sqrt[4]{\underbrace{\sum_{i=1}^{N}\frac{x_{ij}^4}{N^{2-\theta}}\sum_{i=1}^{N}\frac{x_{ik}^4}{N^{2-\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0\,(\text{C1c})}}\sqrt{\underbrace{\sum_{i=1}^{N}\frac{x_{ip}^2\hat{\varepsilon}_i^2}{N}}_{\substack{\text{almost surely}\\ \text{bounded (C1a)}}}}\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0$$

$$\frac{m(x_{ij}^2x_{ik}^2x_{ip}^2\hat{\varepsilon}_i^2/M)}{N^{1-\theta}}=\sum_{i=1}^{N}\frac{x_{ij}^2x_{ik}^2x_{ip}^2\hat{\varepsilon}_i^2}{N^{3-2\theta}}\frac{N}{MN^{\theta}}\le\underbrace{\frac{N}{MN^{\theta}}}_{\rightarrow 0}\sqrt[4]{\underbrace{\sum_{i=1}^{N}\frac{x_{ij}^8}{N^{4-3\theta}}\sum_{i=1}^{N}\frac{x_{ik}^8}{N^{4-3\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0\,(\text{C1d})}}\sqrt{\underbrace{\sum_{i=1}^{N}\frac{x_{ip}^4\hat{\varepsilon}_i^4}{N^{2-\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0\,(\text{D13})}}\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0,$$

so Condition IIIa is met

$$\text{(D17)}\ \underbrace{\frac{m\left(\dfrac{x_{ij}^2x_{ik}^2x_{ip}^2\hat{\varepsilon}_i^2}{M}\right)-m\left(\dfrac{x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_i}{\sqrt{M}}\right)^2}{N^{1-\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}\ 0\ (\text{D16})}\left(\underbrace{\frac{N}{MN^{\theta}}\frac{N}{M}m(\delta_i^2)}_{\overset{p(\boldsymbol{\delta})}{\rightarrow}0\,(\text{L2})}-\underbrace{\frac{N^2}{M^2N^{\theta}}m(\delta_i)^2}_{\rightarrow 0\,(\text{L1})}\right)\overset{p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0.$$

and so

$$\text{(D18) "}e\text{": } m\left(\frac{N}{M^{3/2}}x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_i d_i\right)-\underbrace{m\left(\frac{x_{ij}x_{ik}x_{ip}\hat{\varepsilon}_i}{\sqrt{M}}\right)}_{\text{by (D16)}\ \overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}\ 0}\underbrace{\frac{N}{M}m(\delta_i)}_{=1\,(\text{L1})}\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\rightarrow}0.$$

For "*f*", we apply Theorem III with $z_i = x_{ij}x_{ik}x_{ip}x_{iq}/M$ and "d_i" = $d_i(N/M)$ and note that

$$\text{(D19)}\;\; m\left(\frac{x_{ij}x_{ik}x_{ip}x_{iq}}{M}\right)=\sum_{i=1}^{N}\frac{x_{ij}x_{ik}x_{ip}x_{iq}}{N^{2-\theta}}\frac{N}{MN^{\theta}}\leq\underbrace{\frac{N}{MN^{\theta}}}_{\to 0}\underbrace{\sqrt[4]{\sum_{i=1}^{N}\frac{x_{ij}^4}{N^{2-\theta}}\sum_{i=1}^{N}\frac{x_{ik}^4}{N^{2-\theta}}\sum_{i=1}^{N}\frac{x_{ip}^4}{N^{2-\theta}}\sum_{i=1}^{N}\frac{x_{iq}^4}{N^{2-\theta}}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0\ \text{(C1c)}}\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0$$

$$\frac{m(x_{ij}^2x_{ik}^2x_{ip}^2x_{iq}^2/M^2)}{N^{1-\theta}}=\sum_{i=1}^{N}\frac{x_{ij}^2x_{ik}^2x_{ip}^2x_{iq}^2}{N^{4-3\theta}}\left(\frac{N}{MN^{\theta}}\right)^2\leq\underbrace{\left(\frac{N}{MN^{\theta}}\right)^2}_{\to 0}\underbrace{\sqrt[4]{\sum_{i=1}^{N}\frac{x_{ij}^8}{N^{4-3\theta}}\sum_{i=1}^{N}\frac{x_{ik}^8}{N^{4-3\theta}}\sum_{i=1}^{N}\frac{x_{ip}^8}{N^{4-3\theta}}\sum_{i=1}^{N}\frac{x_{iq}^8}{N^{4-3\theta}}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0\ \text{(C1d)}}\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0$$

and so Condition IIIa is met

$$\text{(D20)}\;\; \underbrace{\frac{m\left(\frac{x_{ij}^2x_{ik}^2x_{ip}^2x_{iq}^2}{M^2}\right)-m\left(\frac{x_{ij}x_{ik}x_{ip}x_{i1}}{M}\right)^2}{N^{1-\theta}}}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0\ \text{(D19)}}\left(\underbrace{\frac{N}{MN^{\theta}}\frac{N}{M}m(\delta_i^2)}_{\overset{p(\boldsymbol{\delta})}{\to} 0\text{(L2)}}-\underbrace{\frac{N^2}{M^2N^{\theta}}m(\delta_i)^2}_{\to 0\text{(L1)}}\right)\overset{p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0$$

and

$$\text{(D21)}\;\; "f":\; m\left(\frac{N}{M^2}x_{ij}x_{ik}x_{ip}x_{iq}d_i\right)-\underbrace{m\left(\frac{x_{ij}x_{ik}x_{ip}x_{iq}}{M}\right)}_{\overset{a.s.(\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0\ \text{(D19)}}\underbrace{\frac{N}{M}m(\delta_i)}_{=1\text{(L1)}}\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\to} 0.$$

Consequently, in (D12) in probability the $\hat{\eta}_p$, which from (D10) converge to random variables with bounded variance, are multiplied by γ_N which converges to 1 and by "*e*" and "*f*" terms which converge to zero, and hence when so multiplied converge in probability to zero. Using (D7), we then establish (D2b):

$$\text{(D22)}\;\; \mathbf{B}-\frac{\mathbf{X}'\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{D}_{\hat{\boldsymbol{\varepsilon}}}\mathbf{X}}{N}\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\to}\mathbf{0}_{KxK}\quad\&\quad M\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_{\mathbf{d},N}^{p})-N\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_N)\overset{p(\mathbf{d})|p(\boldsymbol{\delta},\mathbf{X},\boldsymbol{\varepsilon})}{\to}\mathbf{0}_{KxK}.$$

<u>Proof of the Lemma</u>

We prove the Lemma used above. (L4) will hold automatically, as the mean of δ_i will always equal $M/N < 1$ and the δ_i only take on integer values. For the case of sampling without replacement, δ_i is 1 for M observations and 0 for N-M, and so we have

$$\text{(L1) } m(\delta_i)=\sum_{i=1}^{N}\frac{\delta_i}{N}=\frac{M}{N};\quad \text{(L2)}\left(\frac{N}{MN^{\theta}}\right)\frac{N}{M}m(\delta_i^2)=\frac{N}{MN^{\theta}}=O(N^{1-\theta-\gamma^*})\to 0;$$

$$\text{(L3)}\ \frac{\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^{\tau}}{N^{1+\theta\left(\frac{\tau}{2}-1\right)}}}{\left(\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^2}{N}\right)^{\tau/2}}=\frac{\left(-\frac{M}{N}\right)^{\tau}+\sum_{v=1}^{\tau}\frac{\tau!}{v!(\tau-v)!}\left(\frac{M}{N}\right)\left(-\frac{M}{N}\right)^{\tau-v}}{N^{\theta\left(\frac{\tau}{2}-1\right)}\left(\frac{M}{N}\left(1-\frac{M}{N}\right)\right)^{\tau/2}}=$$

$$\frac{\left(\frac{N}{MN^{\theta}}\right)^{\tau/2-1}\left[\left(-\frac{M}{N}\right)^{\tau}\frac{N}{M}+\sum_{v=1}^{\tau}\frac{\tau!}{v!(\tau-v)!}\left(-\frac{M}{N}\right)^{\tau-v}\right]}{\left(1-\frac{M}{N}\right)^{\tau/2}}=O(N^{(1-\theta-\gamma^*)(\tau/2-1)})\to 0;$$

$$\text{(L5)}\sum_{i=1}^{N}\frac{[\delta_i-m(\delta_i)]^2}{M}=1-\frac{M}{N}\to 1.$$

For sampling with replacement, we define the random variable c_{ti} as in Appendix B above and following the notation and methods used there see that:

$$\text{(D23) } E\left(\frac{N}{M}m(\delta_i)\right)=\frac{1}{M}\sum_{i=1}^{N}\sum_{t=1}^{M}E(c_{ti})=\frac{1}{M}\sum_{i=1}^{N}\sum_{t=1}^{M}\frac{1}{N}=1,$$

$$E\left(\frac{N^2}{M^2}m(\delta_i)^2\right)=E\left[\frac{1}{M^2}\sum_{i=1}^{N}\sum_{t=1}^{M}\sum_{j=1}^{N}\sum_{s=1}^{M}c_{ti}c_{sj}\right]=\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{M}E(c_{ti}c_{tj})+\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s,t=1}^{M}E(c_{ti}c_{sj})$$

$$=\frac{1}{M^2}\sum_{i=1}^{N}\sum_{t=1}^{M}\frac{1}{N}+\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{s,t=1}^{M}\frac{1}{N^2}=\frac{1}{M}+\frac{M(M-1)}{M^2}=1,$$

$$\&\ E\left(\frac{N^2}{M^2}m(\delta_i)^2\right)-E\left(\frac{N}{M}m(\delta_i)\right)^2=0\ \text{[as expected]},$$

So, $m(\delta_i) = M/N$ is a constant with zero variance (proving L1 in the Lemma used above).

Continuing with the notation and techniques of Appendix B above, we calculate the expectation of $(N/M)m(\delta_i^{\tau})$ for integer $\tau \geq 2$.

$$\text{(D24) } E\left(\frac{N}{M}m(\delta_i^{\tau})\right)=\frac{1}{M}\sum_{i=1}^{N}\sum_{t_1=1}^{M}...\sum_{t_\tau=1}^{M}E(c_{t_1 i}...c_{t_\tau i})=\frac{N}{M}\sum_{t_1=1}^{M}...\sum_{t_\tau=1}^{M}E(c_{t_1 i}...c_{t_\tau i})=\frac{N}{M}c_N(\tau)$$

$$\text{with } c_N(\tau)=\sum_{j=1}^{\tau}\frac{M!/(M-j)!}{N^j}C_j^{\tau}$$

and the expectation of $(N/M)^2 m(\delta_i^{\tau})^2$

$$\text{(D25) } E\left(\frac{N^2}{M^2}m(\delta_i^\tau)^2\right)=\frac{1}{M^2}\sum_{i=1}^{N}\sum_{t_1=1}^{M}\cdots\sum_{t_\tau=1}^{M}\sum_{j=1}^{N}\sum_{s_1=1}^{M}\cdots\sum_{s_\tau=1}^{M}E(c_{t_1 i}\ldots c_{t_\tau i}c_{s_1 j}\ldots c_{s_\tau j})$$

$$=\frac{1}{M^2}\sum_{i=1}^{N}\sum_{t_1=1}^{M}\cdots\sum_{t_\tau=1}^{M}\sum_{s_1=1}^{M}\cdots\sum_{s_\tau=1}^{M}E(c_{t_1 i}\ldots c_{t_\tau i}c_{s_1 i}\ldots c_{s_\tau i})+\frac{1}{M^2}\sum_{i,j=1}^{N}\sum_{t_1=1}^{M}\cdots\sum_{t_\tau=1}^{M}\sum_{s_1=1}^{M}\cdots\sum_{s_\tau=1}^{M}E(c_{t_1 i}\ldots c_{t_\tau i}c_{s_1 j}\ldots c_{s_\tau j})$$

$$=\frac{N}{M^2}\underbrace{\sum_{t_1=1}^{M}\cdots\sum_{t_\tau=1}^{M}\sum_{s_1=1}^{M}\cdots\sum_{s_\tau=1}^{M}E(c_{t_1 i}\ldots c_{t_\tau i}c_{s_1 i}\ldots c_{s_\tau i})}_{d_N(\tau)}+\frac{N(N-1)}{M^2}\underbrace{\sum_{t_1=1}^{M}\cdots\sum_{t_\tau=1}^{M}\sum_{s_1=1}^{M}\cdots\sum_{s_\tau=1}^{M}E(c_{t_1 i}\ldots c_{t_\tau i}c_{s_1 j}\ldots c_{s_\tau j})}_{e_N(\tau)}$$

$$\text{with } d_N(\tau)=\sum_{j=1}^{2\tau}\frac{M!/(M-j)!}{N^j}C_j^{2\tau} \text{ and } e_N(\tau)=\sum_{j=1}^{\tau}\sum_{k=1}^{\tau}\frac{M!/(M-j-k)!}{N^{j+k}}C_j^\tau C_k^\tau.$$

We note that

$$\text{(D26) } e_N(\tau)-c_N(\tau)^2=\sum_{j=1}^{\tau}\sum_{k=1}^{\tau}\frac{M!}{N^{j+k}(M-j-k)!}C_j^\tau C_k^\tau-\sum_{j=1}^{\tau}\sum_{k=1}^{\tau}\frac{M!M!}{N^{j+k}(M-j)!(M-k)!}C_j^\tau C_k^\tau$$

$$=\sum_{j=1}^{\tau}\sum_{k=1}^{\tau}\frac{M!C_j^\tau C_k^\tau}{N^{j+k}(M-j-k)!}\left[1-\frac{M(M-1)\ldots(M-j+1)}{(M-k)(M-k+1)\ldots(M-k-j+1)}\right]<0.$$

From this we see that $(N/M)m(\delta_i^\tau)$ converges in mean square to the finite value given in (D24) as its variance is $O(M^{-1})$

$$\text{(D27) } E\left[\frac{N}{M}m(\delta_i^\tau)-E\left[\frac{N}{M}m(\delta_i^\tau)\right]\right]^2=E\left[\frac{N^2}{M^2}m(\delta_i^\tau)^2\right]-E\left[\frac{N}{M}m(\delta_i^\tau)\right]^2$$

$$=\frac{Nd_N(\tau)}{M^2}+\frac{N(N-1)}{M^2}e_N(\tau)-\frac{N^2}{M^2}c_N(\tau)^2<\frac{Nd_N(\tau)}{M^2}$$

$$=\frac{1}{M}\sum_{j=1}^{2\tau}\frac{M-1!/(M-j)!}{N^{j-1}}C_j^{2\tau}<\frac{1}{M}\sum_{j=1}^{2\tau}C_j^{2\tau}.$$

Since $N/MN^\theta\to 0$, this establishes (L2).

For $\tau=2$, we have

$$\text{(D28) } E\left(\frac{N}{M}m(\delta_i^2)\right)=\frac{N}{M}c_N(\tau)=\sum_{j=1}^{2}\frac{M!/(M-j)!}{MN^{j-1}}C_j^2=1+\frac{M-1}{N}$$

$$\text{so } \sum_{i=1}^{N}\frac{N}{M}\frac{[\delta_i-m(\delta_i)]^2}{N}=\frac{N}{M}m(\delta_i^2)-\frac{N}{M}m(\delta_i)^2\xrightarrow{p(\boldsymbol{\delta})}1+\frac{M-1}{N}-\frac{M}{N}\to 1,$$

which establishes (L5). For $\tau>2$

$$\text{(D29)}\; E\left(\sum_{i=1}^{N}\frac{N}{M}\frac{[\delta_i - m(\delta_i)]^\tau}{N}\right) = \sum_{v=0}^{\tau}\frac{\tau!}{v!(\tau-v)!}E\left(\frac{N}{M}m(\delta_i^v)\right)\left(-\frac{M}{N}\right)^{\tau-v}$$

$$\overset{p(\boldsymbol{\delta})}{\rightarrow}\left(-\frac{M}{N}\right)^{\tau}\frac{N}{M} + \sum_{v=1}^{\tau}\frac{\tau!}{v!(\tau-v)!}\left(\sum_{j=1}^{v}\frac{M!/(M-j)!}{MN^{j-1}}C_j^v\right)\left(-\frac{M}{N}\right)^{\tau-v}$$

which is a finite constant. Consequently,

$$\text{(D30)}\; \frac{\sum_{i=1}^{N}\frac{[\delta_i - m(\delta_i)]^\tau}{N^{1+\theta\left(\frac{\tau}{2}-1\right)}}}{\left(\sum_{i=1}^{N}\frac{[\delta_i - m(\delta_i)]^2}{N}\right)^{\tau/2}} = \frac{\left(\frac{N}{MN^\theta}\right)^{\frac{\tau}{2}-1}\sum_{i=1}^{N}\frac{N}{M}\frac{[\delta_i - m(\delta_i)]^\tau}{N}}{\left(\sum_{i=1}^{N}\frac{N}{M}\frac{[\delta_i - m(\delta_i)]^2}{N}\right)^{\tau/2}} \overset{p(\boldsymbol{\delta})}{\rightarrow} 0,$$

which establishes (L3) and completes the proof.