Supplementary Appendix

Parameter-invariant unbiased estimation of individual variances and their pairwise products

Stanislav Anatolyev* — CERGE-EI and NES

Yaroslav Korobka† — CERGE-EI

Contents

SA1 Redefinitions and unbiasedness **2**

SA2 Optimal sample splitting **4**

SA3 Algorithm for efficient four-splitting **6**

SA4 Parameter-invariant test for many restrictions **8**

SA5 More simulation evidence **11**

*Corresponding author. Address: CERGE-EI, Politických vězňů 7, 11121 Prague 1, Czech Republic. E-mail: stanislav.anatolyev@cerge-ei.cz.

†Address: CERGE-EI, Politických vězňů 7, 11121 Prague 1, Czech Republic. E-mail: yaroslav.korobka@cerge-ei.cz.

SA1 Redefinitions and unbiasedness

Henceforth, we denote the regressor matrix by X. Define the full-sample annihilation matrix

$$M = I_n - X\left(X'X\right)^{-1}X',$$

with i^{th} diagonal element M_{ii}. Leave-one-out algebra implies

$$\hat{e}_{-i} = \frac{1}{M_{ii}}\sum_{j\in\mathcal{H}} M_{ij}e_j.$$

Therefore, for the KSS estimator $\hat{\sigma}_i^2$,

$$\begin{aligned} E\left[\hat{\sigma}_i^2|X\right] &= \frac{1}{M_{ii}}E\left[\Big(\sum_{j\in\mathcal{H}} M_{ij}e_j\Big)\left(x_i'\theta + e_i\right)|X\right] \\ &= M_{ii}^{-1}\sum_{j\in\mathcal{H}} M_{ij}E\left[e_ie_j|X\right] \\ &= \sigma_i^2. \end{aligned}$$

Thus, the variance estimator $\hat{\sigma}_i^2$ is conditionally unbiased.

Define the split-sample annihilation matrices

$$M_{\mathcal{H}_i} = I_{|\mathcal{H}_i|} - X_{\mathcal{H}_i}\left(X'_{\mathcal{H}_i}X_{\mathcal{H}_i}\right)^{-1}X'_{\mathcal{H}_i}$$

based only on regressors $X_{\mathcal{H}_i}$ included in $\mathcal{H}_i$, with $(i,j)^{th}$ element $M_{\mathcal{H}_i,ij}$. Leave-one-out algebra implies

$$\mathring{e}_{\mathcal{H}_i,-i} = \frac{1}{M_{\mathcal{H}_i,ii}}\sum_{j\in\mathcal{H}_i} M_{\mathcal{H}_i,ij}e_j.$$

Denote, for $i\in\mathcal{H}_i$ and $k\in\mathcal{H}_{-i}$,

$$M_{\mathcal{H}_{-i},ik} = -x_i'\big(X'_{\mathcal{H}_{-i}}X_{\mathcal{H}_{-i}}\big)^{-1}x_k.$$

Then, for the CF estimator $\mathring{\sigma}_i^2$,

$$\begin{aligned} E\left[\mathring{\sigma}_i^2|X\right] &= \frac{1}{M_{\mathcal{H}_i,ii}}E\left[\Big(\sum_{j\in\mathcal{H}_i} M_{\mathcal{H}_i,ij}e_j\Big)\Big(e_i + \sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\Big)|X\right] \\ &= M_{\mathcal{H}_i,ii}^{-1}\sum_{j\in\mathcal{H}_i} M_{\mathcal{H}_i,ij}E\left[e_je_i|X\right] + M_{\mathcal{H}_i,ii}^{-1}\sum_{j\in\mathcal{H}_i}\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_i,ij}M_{\mathcal{H}_{-i},ik}E\left[e_je_k|X\right] \\ &= M_{\mathcal{H}_i,ii}^{-1}M_{\mathcal{H}_i,ii}E\left[e_i^2|X\right] + 0 \\ &= \sigma_i^2. \end{aligned}$$

Thus, the variance estimator $\mathring{\sigma}_i^2$ is conditionally unbiased.

For the CF variance product estimator $\mathring{\omega}_{ij}$, similar notation for annihilator matrices is extended to subsamples $\mathcal{H}_{(i,-j)}$, $\mathcal{H}_{(-i,j)}$, $\mathcal{H}^1_{(-i,-j)}$ and $\mathcal{H}^2_{(-i,-j)}$, together with the leave-one-out and regular OLS residual identities. Then, the expectation $\mathbb{E}[M_{\mathcal{H}_{(i,-j)},ii}M_{\mathcal{H}_{(-i,j)},jj}\mathring{\omega}_{ij}|X]$ equals

$$\begin{aligned}
& \mathbb{E}\Bigg[\sum_{k\in\mathcal{H}_{(i,-j)}} M_{\mathcal{H}_{(i,-j)},ik}e_k \sum_{p\in\mathcal{H}_{(-i,j)}} M_{\mathcal{H}_{(-i,j)},jp}e_p \\
&\qquad \times\Big(e_i + \sum_{q\in\mathcal{H}^1_{(-i,-j)}} M_{\mathcal{H}^1_{(-i,-j)},iq}e_q\Big)\Big(e_j + \sum_{r\in\mathcal{H}^2_{(-i,-j)}} M_{\mathcal{H}^2_{(-i,-j)},jr}e_r\Big)|X\Bigg] \\
= & \sum_{k\in\mathcal{H}_{(i,-j)}}\sum_{p\in\mathcal{H}_{(-i,j)}} M_{\mathcal{H}_{(i,-j)},ik}M_{\mathcal{H}_{(-i,j)},jp}\mathbb{E}[e_ie_je_ke_p|X] \\
& + \sum_{k\in\mathcal{H}_{(i,-j)}}\sum_{p\in\mathcal{H}_{(-i,j)}}\sum_{r\in\mathcal{H}^2_{(-i,-j)}} M_{\mathcal{H}_{(i,-j)},ik}M_{\mathcal{H}_{(-i,j)},jp}M_{\mathcal{H}_{(-i,-j)},jr}\mathbb{E}[e_ie_ke_pe_r|X] \\
& + \sum_{k\in\mathcal{H}_{(i,-j)}}\sum_{p\in\mathcal{H}_{(-i,j)}}\sum_{q\in\mathcal{H}^1_{(-i,-j)}} M_{\mathcal{H}_{(i,-j)},ik}M_{\mathcal{H}_{(-i,j)},jp}M_{\mathcal{H}_{(-i,-j)},iq}\mathbb{E}[e_je_ke_pe_q|X] \\
& + \sum_{k\in\mathcal{H}_{(i,-j)}}\sum_{p\in\mathcal{H}_{(-i,j)}}\sum_{q\in\mathcal{H}^1_{(-i,-j)}}\sum_{r\in\mathcal{H}^2_{(-i,-j)}} M_{\mathcal{H}_{(i,-j)},ik}M_{\mathcal{H}_{(-i,j)},jp}M_{\mathcal{H}_{(-i,-j)},iq}M_{\mathcal{H}_{(-i,-j)},jr}\mathbb{E}[e_ke_pe_qe_r|X] \\
= & M_{\mathcal{H}_{(i,-j)},ii}M_{\mathcal{H}_{(-i,j)},jj}\mathbb{E}[e_i^2e_j^2|X] + 0 + 0 + 0 \\
= & M_{\mathcal{H}_{(i,-j)},ii}M_{\mathcal{H}_{(-i,j)},jj}\omega_{ij},
\end{aligned}$$

so $\mathbb{E}[\mathring{\omega}_{ij}|X] = \omega_{ij}$. Thus, the variance product estimator $\mathring{\omega}_{ij}$ is conditionally unbiased.

SA2 Optimal sample splitting

Towards the goal of finding the optimal split to the index sets $\mathcal{H}_i$ and $\mathcal{H}_{-i}$, compute the conditional variance of $\mathring{\sigma}_i^2$:

$$\begin{aligned}
M^2_{\mathcal{H}_i,ii} var\left(\mathring{\sigma}_i^2|X\right) &= \mathbb{E}\Bigg[\Big(M_{\mathcal{H}_i,ii}e_i + \sum_{\substack{j\in\mathcal{H}_i\\ j\neq i}} M_{\mathcal{H}_i,ij}e_j\Big)^2\Big(e_i + \sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\Big)^2|X\Bigg] - M^2_{\mathcal{H}_i,ii}\sigma_i^4 \\
&= \mathbb{E}\left[\begin{array}{c}\Big(M^2_{\mathcal{H}_i,ii}e_i^2 + \big(\sum_{\substack{j\in\mathcal{H}_i\\ j\neq i}} M_{\mathcal{H}_i,ij}e_j\big)^2 + 2M_{\mathcal{H}_i,ii}e_i\sum_{\substack{j\in\mathcal{H}_i\\ j\neq i}} M_{\mathcal{H}_i,ij}e_j\Big) \\ \times\Big(e_i^2 + \big(\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big)^2 + 2e_i\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\Big)|X\end{array}\right] \\
&\quad - M^2_{\mathcal{H}_i,ii}\sigma_i^4 \\
&= \mathbb{E}\left[\begin{array}{l} M^2_{\mathcal{H}_i,ii}e_i^4 + \big(\sum_{j\in\mathcal{H}_i,j\neq i} M_{\mathcal{H}_i,ij}e_j\big)^2\big(\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big)^2 \\ +4M_{\mathcal{H}_i,ii}e_i^2\big(\sum_{j\in\mathcal{H}_i,j\neq i} M_{\mathcal{H}_i,ij}e_j\big)\big(\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big) \\ +e_i^2\big(\sum_{j\in\mathcal{H}_i,j\neq i} M_{\mathcal{H}_i,ij}e_j\big)^2 + M^2_{\mathcal{H}_i,ii}e_i^2\big(\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big)^2 \\ +2e_i^3\big(M_{\mathcal{H}_i,ii}\sum_{j\in\mathcal{H}_i,j\neq i} M_{\mathcal{H}_i,ij}e_j + M^2_{\mathcal{H}_i,ii}\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big) \\ +2e_i\big(\sum_{j\in\mathcal{H}_i,j\neq i} M_{\mathcal{H}_i,ij}e_j\big)^2\big(\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big) \\ +2M_{\mathcal{H}_i,ii}e_i\big(\sum_{j\in\mathcal{H}_i,j\neq i} M_{\mathcal{H}_i,ij}e_j\big)\big(\sum_{k\in\mathcal{H}_{-i}} M_{\mathcal{H}_{-i},ik}e_k\big)^2|X \end{array}\right] \\
&\quad - M^2_{\mathcal{H}_i,ii}\sigma_i^4 \\
&= M^2_{\mathcal{H}_i,ii}\left(\kappa_i - \sigma_i^4\right) + \Big(\sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij}\sigma_j^2\Big)\Big(\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}\sigma_k^2\Big) \\
&\quad + \sigma_i^2\Big(\sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij}\sigma_j^2 + M^2_{\mathcal{H}_i,ii}\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}\sigma_k^2\Big),
\end{aligned}$$

exploiting conditional independence across $\mathcal{H}_i$ and $\mathcal{H}_{-i}$. Here, $\kappa_i = \mathbb{E}[e_i^4|X]$. Then,

$$\begin{aligned}
var\left(\mathring{\sigma}_i^2|X\right) &= \kappa_i - \sigma_i^4 + \frac{1}{M^2_{\mathcal{H}_i,ii}}\Big(\sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij}\sigma_j^2\Big)\Big(\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}\sigma_k^2\Big) \\
&\quad + \frac{\sigma_i^2}{M^2_{\mathcal{H}_i,ii}}\Big(\sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij}\sigma_j^2 + M^2_{\mathcal{H}_i,ii}\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}\sigma_k^2\Big).
\end{aligned}$$

Suppose heteroskedasticity is bounded: $\underline{\sigma}^2 \leq \sigma_j^2 \leq \bar{\sigma}^2$ for all $j \in \mathcal{H}$. Then,

$$\underline{\sigma}^2\sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij} \leq \sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij}\sigma_j^2 \leq \bar{\sigma}^2\sum_{j\in\mathcal{H}_i,j\neq i} M^2_{\mathcal{H}_i,ij}$$

and similarly,

$$\underline{\sigma}^2\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik} \leq \sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}\sigma_k^2 \leq \bar{\sigma}^2\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}.$$

Further note that $\sum_{j\in\mathcal{H}_i, j\neq i} M^2_{\mathcal{H}_i,ij} = M_{\mathcal{H}_i,ii}\,(1-M_{\mathcal{H}_i,ii})$ and $\sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik} = 1 - M_{\mathcal{H}_{-i},ii}$.

When $m = o\big(\min\{|\mathcal{H}_i|, |\mathcal{H}_{-i}|\}\big)$, we have $M_{\mathcal{H}_i,ii} = 1+o(1)$, $M_{\mathcal{H}_i,ii}\,(1-M_{\mathcal{H}_i,ii}) = o(1)$, $1-M_{\mathcal{H}_{-i},ii} = o\,(1)$, and hence

$$var\left(\mathring{\sigma}_i^2|X\right) - \left(\kappa_i - \sigma_i^4\right) = \sigma_i^2\Big(\sum_{j\in\mathcal{H}_i} M^2_{\mathcal{H}_i,ij}\sigma_j^2 + \sum_{k\in\mathcal{H}_{-i}} M^2_{\mathcal{H}_{-i},ik}\sigma_k^2\Big)\left(1+o(1)\right),$$

which, apart from the multiplicative factor that asymptotically does not depend on the sample split, equals

$$(1-M_{\mathcal{H}_i,ii}) + \left(1-M_{\mathcal{H}_{-i},ii}\right).$$

This quantity depends on the whole regressor set. Averaged over corresponding subsamples, however, this quantity equals

$$\frac{1}{|\mathcal{H}_i|}\sum_{i\in\mathcal{H}_i}(1-M_{\mathcal{H}_i,ii}) + \frac{1}{|\mathcal{H}_{-i}|}\sum_{i\in\mathcal{H}_{-i}}\left(1-M_{\mathcal{H}_{-i},ii}\right) = \frac{m}{|\mathcal{H}|}\left(\frac{1}{\eta_i} + \frac{1}{1-\eta_i}\right),$$

where $\eta_i = |\mathcal{H}_i|\,/\,|\mathcal{H}|$, and is minimized on $[0,1]$ at $\eta_i = \frac{1}{2}$. The same result obtains if instead one relies on unconditional expectations, in which case, asymptotically, $\mathbb{E}\left[1-M_{\mathcal{H}_i,ii}\right] \asymp m/\,|\mathcal{H}_i|$ and $\mathbb{E}\left[1-M_{\mathcal{H}_{-i},ii}\right] \asymp m/\,|\mathcal{H}_{-i}|$.

SA3 Algorithm for efficient four-splitting

The four-split estimator $\mathring{\omega}_{ij}$ requires four random subsamples with an additional restriction placed on which subsamples the indices i and j belong to. Specifically, splitting the full sample randomly into four subsamples may produce a subsample that contains both i^{th} and j^{th} observations, which violates the sample split restriction $\mathcal{H} = \mathcal{H}_{(i,-j)} \cup \mathcal{H}_{(-i,j)} \cup \mathcal{H}^1_{(-i,-j)} \cup \mathcal{H}^2_{(-i,-j)}$. Below, we describe a simple algorithm that always produces random subsamples satisfying this restriction. The algorithm is efficient in a sense that for all pairs (i, j), it requires only one split, and computes an inverse matrix for each subsample only once. When observations i and j belong to the same subsample, we randomly swap observation j with observation q from another subsample, and recompute the matrices using the Woodbury rank-2 update. This strategy avoids expensive computation of inverses separately for each (i, j).

1. Randomly split $\mathcal{H}$ into sets $\mathcal{H}_k$ for $k = 1, \ldots, 4$, each with size $\lfloor n/4 \rfloor$.

2. For each $k = 1, \ldots, 4$, compute $\tilde{X}_k := \left(\sum_{i \in \mathcal{H}_k} x_i x_i'\right)^{-1}$. Then...

   (i) ...if the collection of sets $\{\mathcal{H}_k\}_{k=1}^4$ satisfies the sample split restriction, rename the sets as $\mathcal{H}_{(i,-j)}$,$\mathcal{H}_{(-i,j)}$, $\mathcal{H}^1_{(-i,-j)}$, $\mathcal{H}^2_{(-i,-j)}$ depending on whether i or j belongs to the set. Without a loss of generality, let $i \in \mathcal{H}_1$ and $j \notin \mathcal{H}_1$, $i \notin \mathcal{H}_2$ and $j \in \mathcal{H}_2$, and $i, j \notin \mathcal{H}_r$ for $r \in \{3, 4\}$. Then, set

   $$\mathcal{H}_{(i,-j)} := \mathcal{H}_1, \quad \mathcal{H}_{(-i,j)} := \mathcal{H}_2, \quad \mathcal{H}^1_{(-i,-j)} := \mathcal{H}_3, \quad \mathcal{H}^2_{(-i,-j)} := \mathcal{H}_4.$$

   Construct residuals corresponding to these sets.

   (ii) ...if for some r, we have that $i, j \in \mathcal{H}_r$, swap j with any randomly selected observation q from another set $\mathcal{H}_p$, and rename the sets as $\mathcal{H}_{(i,-j)}$,$\mathcal{H}_{(-i,j)}$, $\mathcal{H}^1_{(-i,-j)}$, $\mathcal{H}^2_{(-i,-j)}$ depending on whether i or j belongs to the set as in (i). Using the Woodbury rank-2 update, recompute $\tilde{X}_k$ for $k \in \{r, p\}$ as

   $$\tilde{X}_{(r,-j)} := \tilde{X}_r + \frac{\tilde{X}_r x_j x_j' \tilde{X}_r}{1 - x_j' \tilde{X}_r x_j}, \quad \tilde{X}_{(r,-j,q)} := \tilde{X}_{(r,-j)} - \frac{\tilde{X}_{(r,-j)} x_q x_q' \tilde{X}_{(r,-j)}}{1 + x_q' \tilde{X}_{(r,-j)} x_q},$$
   $$\tilde{X}_{(p,-q)} := \tilde{X}_p + \frac{\tilde{X}_p x_q x_q' \tilde{X}_p}{1 - x_q' \tilde{X}_p x_q}, \quad \tilde{X}_{(p,-q,j)} := \tilde{X}_{(p,-q)} - \frac{\tilde{X}_{(p,-q)} x_j x_j' \tilde{X}_{(p,-q)}}{1 + x_j' \tilde{X}_{(p,-q)} x_j},$$

and update

$$\check{y}_{(r,-j,q)} := \sum_{i \in \mathcal{H}_r, i \neq j} x_i y_i + x_q y_q, \quad \check{y}_{(p,-q,j)} := \sum_{i \in \mathcal{H}_p, i \neq q} x_i y_i + x_j y_j.$$

Now, use $(\check{X}_{(r,-j,q)}, \check{y}_{(r,-j,q)})$ to construct residuals corresponding to the set $\mathcal{H}_{(i,-j)}$, use $(\check{X}_{(p,-q,j)}, \check{y}_{(p,-q,j)})$ to construct residuals corresponding to the set $\mathcal{H}_{(-i,j)}$, and use the remaining two unaltered sets to construct residuals corresponding to sets $\mathcal{H}^1_{(-i,-j)}$ and $\mathcal{H}^2_{(-i,-j)}$.

3. If desired, construct the estimator $\ell > 1$ times by repeating steps 1 and 2, and take an average.

SA4 Parameter-invariant test for many restrictions

The test statistic for many restrictions of Anatolyev and Sølvsten (2023) has the form

$$\frac{\mathcal{F}-\hat{E}_{\mathcal{F}}}{\hat{V}_{\mathcal{F}}^{1/2}},$$

which is based on the conventional F-statistic $\mathcal{F}$, an unbiased estimate $\hat{E}_{\mathcal{F}}$ of its conditional expectation $\mathbb{E}\left[\mathcal{F}|X\right]=\sum_{i=1}^{n}B_{ii}\sigma_i^2$ for certain observed quantities B_{ii}, and an unbiased estimate $\hat{V}_{\mathcal{F}}$ of the conditional variance of the difference $\mathcal{F}-\hat{E}_{\mathcal{F}}$. Anatolyev and Sølvsten (2023) use the leave-one-out individual variance estimates, previously referred to as KSS, to construct $\hat{E}_{\mathcal{F}}$, and leave-three-out estimates for the pairwise variance products ω_{ij} to construct $\hat{V}_{\mathcal{F}}$. The term $\hat{E}_{\mathcal{F}}$ that uses KSS estimates, is not parameter-invariant, and the expression for $\hat{V}_{\mathcal{F}}$ (see Anatolyev and Sølvsten 2023, formula (12)) contains two convoluted terms, both of which are not parameter-invariant, both explicitly (via dependence of the conditional variance) and implicitly (via dependence of leave-out estimates).

In what follows, we show how to construct the test for many restrictions that is parameter-invariant. The test statistic has the similar form

$$\frac{\mathcal{F}-\mathring{E}_{\mathcal{F}}}{\mathring{V}_{\mathcal{F}}^{1/2}},$$

where both the unbiased estimate $\mathring{E}_{\mathcal{F}}$ of its conditional expectation $\mathbb{E}\left[\mathcal{F}|X\right]$ and the unbiased estimate $\mathring{V}_{\mathcal{F}}$ of the variance of the difference $\mathcal{F}-\mathring{E}_{\mathcal{F}}$ are parameter-invariant. To estimate the conditional expectation $\mathbb{E}\left[\mathcal{F}|X\right]$, one uses the CF estimates $\mathring{\sigma}_i^2$ based on an arbitrary sample split $\mathcal{H}=\mathcal{H}_1\cup\mathcal{H}_2$. Thus,

$$\mathring{E}_{\mathcal{F}}=\sum_{i=1}^{n}B_{ii}\mathring{\sigma}_i^2$$

is conditionally unbiased for $\mathbb{E}\left[\mathcal{F}|X\right]$ and is parameter-invariant. The difference $\mathcal{F}-\mathring{E}_{\mathcal{F}}$ then equals

$$\begin{aligned}
&\sum_{i\in\mathcal{H}}\sum_{j\in\mathcal{H}}B_{ij}e_ie_j-\sum_{i\in\mathcal{H}_1}\frac{B_{ii}}{M_{\mathcal{H}_1,ii}}\Big(\sum_{j\in\mathcal{H}_1}M_{\mathcal{H}_1,ij}e_j\Big)\Big(e_i+\sum_{k\in\mathcal{H}_2}M_{\mathcal{H}_2,ik}e_k\Big)\\
&-\sum_{i\in\mathcal{H}_2}\frac{B_{ii}}{M_{\mathcal{H}_2,ii}}\Big(\sum_{j\in\mathcal{H}_2}M_{\mathcal{H}_2,ij}e_j\Big)\Big(e_i+\sum_{k\in\mathcal{H}_1}M_{\mathcal{H}_1,ik}e_k\Big)\\
=\;&\sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_1}B_{ij}e_ie_j+\sum_{i\in\mathcal{H}_2}\sum_{j\in\mathcal{H}_2}B_{ij}e_ie_j+\sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_2}B_{ij}e_ie_j+\sum_{i\in\mathcal{H}_2}\sum_{j\in\mathcal{H}_1}B_{ij}e_ie_j\\
&-\sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_1}B_{ii}\frac{M_{\mathcal{H}_1,ij}}{M_{\mathcal{H}_1,ii}}e_ie_j-\sum_{i\in\mathcal{H}_2}\sum_{j\in\mathcal{H}_2}B_{ii}\frac{M_{\mathcal{H}_2,ij}}{M_{\mathcal{H}_2,ii}}e_ie_j\\
=\;&\sum_{i\neq j\in\mathcal{H}_1}C_{1,ij}e_ie_j+\sum_{i\neq j\in\mathcal{H}_2}C_{2,ij}e_ie_j+\sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_2}\left(D_{2,ij}+D_{1,ji}\right)e_ie_j,
\end{aligned}$$

where $C_{\varsigma,ij} = B_{ij} - B_{ii} M_{\mathcal{H}_\varsigma,ij}/M_{\mathcal{H}_\varsigma,ii}$ and $D_{\varsigma,ij} = B_{ij} - \sum_{s\in\mathcal{H}_\varsigma} B_{ss} M_{\mathcal{H}_\varsigma,sj} M_{\mathcal{H}_{-\varsigma},si}/M_{\mathcal{H}_\varsigma,ss}$ for $\varsigma \in \{1,2\}$ and $\mathcal{H}_{-\varsigma} = \mathcal{H}\backslash\mathcal{H}_\varsigma$. Consequently, the conditional variance of $\mathcal{F} - \mathring{E}_\mathcal{F}$ equals

$$var\big(\mathcal{F} - \mathring{E}_\mathcal{F}|X\big) = 2\sum_{i\neq j\in\mathcal{H}_1} C_{1,ij}^2\omega_{ij} + 2\sum_{i\neq j\in\mathcal{H}_2} C_{2,ij}^2\omega_{ij} + \sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_2} \left(D_{2,ij} + D_{1,ji}\right)^2 \omega_{ij}.$$

Note that, in contrast to $var\big(\mathcal{F} - \hat{E}_\mathcal{F}|X\big)$ in the AS test, this expression contains only terms related to products of individual variances, and does not involve objects that explicitly depend on the parameter β, unlike the second term in Anatolyev and Sølvsten (2023, formula (9)). The estimation of each product ω_{ij} by $\mathring{\omega}_{ij}$ leads to the unbiased estimate $\mathring{V}_\mathcal{F}$ of $var\big(\mathcal{F} - \mathring{E}_\mathcal{F}|X\big)$, viz.,

$$\mathring{V}_\mathcal{F} = 2\sum_{i\neq j\in\mathcal{H}_1} C_{1,ij}^2\mathring{\omega}_{ij} + 2\sum_{i\neq j\in\mathcal{H}_2} C_{2,ij}^2\mathring{\omega}_{ij} + \sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_2} \left(D_{2,ij} + D_{1,ji}\right)^2 \mathring{\omega}_{ij}.$$

Importantly, all parts of $\mathring{V}_\mathcal{F}$ are parameter-invariant.

Algorithmically, for the first two "within-subsample" terms, split $\mathcal{H}_\varsigma = \mathcal{H}_\varsigma^a \cup \mathcal{H}_\varsigma^b$ such that $i \in \mathcal{H}_\varsigma^a$ and $j \in \mathcal{H}_\varsigma^b$, and arbitrarily split $\mathcal{H}_{-\varsigma} = \mathcal{H}_{-\varsigma}^a \cup \mathcal{H}_{-\varsigma}^b$. Set

$$\mathcal{H}_{(i,-j)} := \mathcal{H}_\varsigma^a, \quad \mathcal{H}_{(-i,j)} := \mathcal{H}_\varsigma^b, \quad \mathcal{H}_{(-i,-j)}^1 := \mathcal{H}_{-\varsigma}^a, \quad \mathcal{H}_{(-i,-j)}^2 := \mathcal{H}_{-\varsigma}^b.$$

Given these sets, estimate the product ω_{ij} by $\mathring{\omega}_{ij}$. For the last "cross-subsample" term, the indexing is such that $i \in \mathcal{H}_{-\varsigma}$ and $j \in \mathcal{H}_\varsigma$ by construction, hence split further as $\mathcal{H}_\varsigma = \mathcal{H}_\varsigma^a \cup \mathcal{H}_\varsigma^b$ and $\mathcal{H}_{-\varsigma} = \mathcal{H}_{-\varsigma}^a \cup \mathcal{H}_{-\varsigma}^b$. Set

$$\mathcal{H}_{(i,-j)} := \mathcal{H}_{-\varsigma}^\gamma \quad \text{if} \quad i \in \mathcal{H}_{-\varsigma}^\gamma, \quad \mathcal{H}_{(-i,j)} := \mathcal{H}_\varsigma^\gamma \quad \text{if} \quad j \in \mathcal{H}_\varsigma^\gamma, \quad \gamma \in \{a,b\},$$

and set $\mathcal{H}_{(-i,-j)}^1$ to be the part of $\mathcal{H}_{-\varsigma}$ not used for $\mathcal{H}_{(i,-j)}$, and set $\mathcal{H}_{(-i,-j)}^2$ to be the part of $\mathcal{H}_\varsigma$ not used for $\mathcal{H}_{(-i,j)}$.

Finally, note that the a- and b-splits of $\mathcal{H}_\varsigma$ and $\mathcal{H}_{-\varsigma}$ are arbitrary given that indices i and j belong to suitable subsamples. Hence, if desired, one may average over these additional splits, similarly to how one takes averages in computation of $\mathring{\sigma}_i^2$ and $\mathring{\omega}_{ij}$. Then the variance estimator is

$$\mathring{V}_\mathcal{F} = \frac{1}{\ell}\sum_{l=1}^{\ell}\Big(2\sum_{i\neq j\in\mathcal{H}_1}\left(C_{1,ij}^2\mathring{\omega}_{ij}\right)_\ell + 2\sum_{i\neq j\in\mathcal{H}_2}\left(C_{2,ij}^2\mathring{\omega}_{ij}\right)_l + \sum_{i\in\mathcal{H}_1}\sum_{j\in\mathcal{H}_2}\left((D_{2,ij} + D_{1,ji})^2\mathring{\omega}_{ij}\right)_l\Big),$$

where the additional index l indicates that computations are performed in the l^{th} split. We conjecture that such averaging leads to improved precision of the estimator. Moreover, similarly to how it happens with estimates of individual variances, the averaging over splits increases the likelihood of getting a positive variance estimate. In the Monte-Carlo experiments with the AS and AS+ tests in Section 5, $\hat{V}_{\mathcal{F}}$ takes negative values in up to 30% of cases, depending on the value of β, while $\mathring{V}_{\mathcal{F}}$ takes positive values almost always if ℓ is set to a sufficiently large number (we use $\ell = 30$).

Computation-wise, the algorithm described in SA3 is extended to the current setting in a straightforward way.

SA5 More simulation evidence

We have also experimented with the individual variance estimators with some non-normal covariate distributions such as Bernoulli, lognormal, and Pareto. The rejection rate patterns stay similar to those reported in Table 2 in the main text, but the proportions of negative definite variance estimators from the use of KSS increases when the covariate distribution is lognormal (to up to 40%) or Pareto (to up to 45%), though demeaning in KSS+ decreases these figures (to about 20% and about 30%, respectively). The corresponding figures in all cases are virtually zero for CJN/CJN+, and zero by construction for CF.